

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO DE CIÊNCIAS DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA

VERÔNICA LUIZA SILVEIRA FIALHO

**DISCRIMINAÇÃO GENÉTICA DE ESPÉCIES DE *PFAFFIA*
SPP. (GINSENG BRASILEIRO) USANDO CÓDIGO DE
BARRAS DE DNA**

VITÓRIA

2017

VERÔNICA LUIZA SILVEIRA FIALHO

**DISCRIMINAÇÃO GENÉTICA DE ESPÉCIES DE
PFAFFIA SPP. (GINSENG BRASILEIRO) USANDO
CÓDIGO DE BARRAS DE DNA**

Dissertação apresentada ao Programa de Pós-Graduação em Biotecnologia do Centro de Ciências da Saúde da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do título de Mestre em Biotecnologia.

Orientador: Prof^a Dr^a Greiciane Gaburro Paneto.

Co-orientador: Prof. Dr. Iuri Drumond Louro.

VITÓRIA

2017

Dados Internacionais de Catalogação-na-publicação (CIP)
(Biblioteca Setorial do Centro de Ciências da Saúde da Universidade Federal do
Espírito Santo, ES, Brasil)

F438d Fialho, Verônica Luiza Silveira, 1990 -
Discriminação genética de espécies de *Pfaffia* spp. (ginseng brasileiro)
usando código de barras de DNA / Verônica Luiza Silveira Fialho – 2017.
59 f. : il.

Orientador: Greiciane Gaburro Paneto.
Coorientador: Iuri Drumond Louro.

Dissertação (Mestrado em Biotecnologia) – Universidade Federal do
Espírito Santo, Centro de Ciências da Saúde.

1. Ginseng (Homeopatia). 2. DNA. 3. Biotecnologia. I. Paneto, Greiciane
Gaburro. II. Louro, Iuri Drumond. III. Universidade Federal do Espírito Santo.
Centro de Ciências da Saúde. IV. Título.

CDU: 61

VERÔNICA LUIZA SILVEIRA FIALHO

**DISCRIMINAÇÃO GENÉTICA DE ESPÉCIES DE
PFAFFIA SPP. (GINSENG BRASILEIRO) USANDO
CÓDIGO DE BARRAS DE DNA**

Dissertação apresentada ao Programa de Pós-Graduação em Biotecnologia do Centro de Ciências da Saúde da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do título de Mestre em Biotecnologia.

Apresentada em 09 de março de 2017.

Prof^a. Dr^a. Greiciane Gaburro Paneto
Instituição: Universidade Federal do
Espírito Santo

Prof. Dr. Iuri Drumond Louro
Instituição: Universidade Federal do
Espírito Santo

Prof^a. Dr^a. Regina Maria Barretto
Cicarelli
Instituição: Universidade Estadual
Paulista

Prof^a. Dr^a. Flávia de Paula
Instituição: Universidade Federal do
Espírito Santo

VITÓRIA

2017

AGRADECIMENTOS

Agradeço primeiramente a Deus, a toda minha família e ao meu noivo. Agradeço a minha orientadora, Greiciane e ao meu coorientador, Iuri e a todos do laboratório NGHM – Núcleo de Genética Humana e Molecular da UFES, que contribuíram muito com esse trabalho.

Agradeço a FAPES pelo financiamento deste projeto, a CAPES pela bolsa concedida e ao programa de Pós-Graduação em Biotecnologia da UFES.

RESUMO

Plantas das espécies *Pfaffia* spp. são amplamente utilizadas em território nacional e são popularmente conhecidas como Ginseng-brasileiro. As espécies mais conhecidas são a *Pfaffia glomerata* e a *Pfaffia paniculata*. Há indicações de uso como tônico e revigorante geral, no tratamento de fadiga física, esgotamento mental, falta de memória, como auxiliar no tratamento de distúrbios circulatórios, dentre outros. Embora as duas espécies sejam utilizadas com os mesmos propósitos, deve-se considerar que a diferença química entre elas pode interferir nos espectros de ações farmacológicas e toxicológicas. A identificação morfológica das raízes é tarefa difícil sobretudo devido a sua forma de comercialização. Por isso, faz-se necessário o desenvolvimento de novas metodologias de identificação das espécies comercializadas. O objetivo principal deste trabalho foi identificar, em nível de espécie, amostras de Ginseng-brasileiro vendidos no mercado brasileiro, utilizando, para isso, o código de barras de DNA (*DNA barcode*). O DNA de 60 amostras comerciais foi extraído e os genes *matK* e *rbcL* foram amplificados por PCR e sequenciados. Além disso, amostras referência, morfológicamente identificadas, das duas espécies foram utilizadas para comparação e submetidas ao mesmo protocolo. Posteriormente, as amostras foram confrontadas com amostras referências e com o banco de dados BOLD (*Barcode of Life Data Systems*). Dentre as amostras referências, 90% amplificaram, tanto para *matK* quanto para *rbcL*. As amostras comerciais obtiveram taxa de amplificação de 95% para ambos os genes. Todas as amostras foram sequenciadas. Encontrou-se uma diferença interespecífica entre *P. glomerata* e *P. paniculata* de 1,92% para *matK* e 1,90% para *rbcL*. Não foi encontrada variação intraespecífica. Das 58 amostras avaliadas, 67,24% foram identificadas, sendo 61,54% *P. glomerata* e 38,46% *P. paniculata*. O restante, 32,76% não foram identificadas conforme o rótulo, caracterizando adulteração. Desses, 22,41% não foram identificadas e 10,34% apresentaram substituição de espécies. O método de identificação por DNA barcode possibilitou a identificação, em nível de espécie, das amostras comercializadas como Ginseng brasileiro obtidas do mercado brasileiro, como *P. glomerata* e *P. paniculata*.

Palavras-chave: Planta medicinal. Ginseng. *Pfaffia*. DNA *Barcoding*.

ABSTRACT

Plants of *Pfaffia* spp. species are widely used in national territory and are popularly known as Ginseng-Brazilian. The best-known species are *Pfaffia glomerata* and *Pfaffia paniculata*. There are indications of use as tonic and general invigorating, in the treatment of physical fatigue, mental exhaustion, lack of memory, as an aid in the treatment of circulatory disorders, among others. Although the two species are used for the same purposes, it must be considered that the chemical difference between them may interfere with the spectra of pharmacological and toxicological actions. The morphological identification of the roots is difficult task, mainly due to its commercialization. Therefore, it is necessary to develop new methodologies for the identification of commercialized species. The main objective of this work was to identify, at species level, Ginseng-Brazilian samples sold in the Brazilian market, using the DNA barcode. DNA from 60 commercial samples was extracted and the *matK* and *rbcl* genes were amplified by PCR and sequenced. In addition, reference samples, morphologically identified, of the two species were used for comparison and submitted to the same protocol. Subsequently, the samples were compared with reference samples and with the BOLD (*Barcode of Life Data Systems*) database. Among the reference samples, 90% amplified for both *matK* and *rbcl* genes. Commercial samples obtained a 95% amplification rate for both genes. All samples were sequenced. An interspecific difference was found between *P. glomerata* and *P. paniculata* of 1.92% for *matK* and 1.90% for *rbcl*. No intraspecific variation was found. Of the 58 samples evaluated, 67.24% were identified, being 61.54% *P. glomerata* and 38.46% *P. paniculata*. The remainder, 32.76% were not identified according to the label, characterizing adulteration. Of these, 22.41% were not identified and 10.34% presented species substitution. The method of identification by DNA barcode allowed the identification, at species level, of samples commercialized as Brazilian Ginseng obtained from the Brazilian market, such as *P. glomerata* and *P. paniculata*.

Keywords: Medicinal plant. Ginseng. *Pfaffia*. DNA *Barcoding*.

LISTA DE FIGURAS

Figura 1 – Resultado de uma busca do gene <i>rbcL</i> da planta <i>Pfaffia</i> <i>glomerata</i> no BOLD.....	18
Figura 2 –Site do banco de dados BOLD.....	22
Figura 3 –Eletroferograma da amostra 60, gene <i>rbcL</i> , indicando picos duplos e indicação de mistura de espécie.....	39
Figura 4 –Eletroferograma da amostra comercializada 31, do gene <i>matK</i> , demonstrando repetição de bases.....	40

LISTA DE TABELAS

Tabela 1 – <i>Primers</i> utilizados no estudo.....	26
Tabela 2 – Reagentes e quantidade utilizados na reação de PCR.....	26
Tabela 3 – Reagentes e quantidades utilizados para reação de sequenciamento....	28
Tabela 4 – Taxas de variações intraespecífica e interespecífica <i>P. glomerata</i> e <i>P. Paniculata</i> encontradas em amostras referência deste trabalho.....	34
Tabela 5 – Identificação das amostras comerciais após confronto com as amostras referências.....	35
Tabela 6 – Similaridade de amostras não identificadas com espécies depositadas no BOLD.....	37

LISTA DE SIGLAS

ANVISA	Agência Nacional de Vigilância Sanitária
BOLD	<i>Barcode of Life Data Systems</i>
BrBOL	<i>Brazilian Barcode of Life</i>
C	Celsius
CBOL	<i>Consortium for the Barcoding of Life</i>
DNA	Ácido desoxirribonucleico (do inglês deoxyribonucleic acid)
EDTA	Ácido etilenodiamino tetra-acético (do inglês <i>Ethylenediamine tetraacetic acid</i>)
EXO I	Exonuclease I
matK	<i>maturase K</i>
mM	Micromolar
MT	Mato Grosso
Ng	Nanogramas
NJ	Neighbor-Joining
OMS	Organização Mundial da Saúde
PCR	Reação em cadeia da polimerase (do inglês <i>Polymerase Chain Reaction</i>)
pM	Picomolar
rbcL	<i>ribulose-bisphosphate carboxylase</i>
SAP	<i>Shrimp Alkaline Phosphatase</i>
SP	São Paulo
U	Unidade
UFMT	Universidade Federal do Mato Grosso
UNICAMP	Universidade Estadual de Campinas
μl	Microlitro

SUMÁRIO

1 INTRODUÇÃO.....	11
1.1 Uso popular de plantas medicinais.....	11
1.2 Uso de plantas medicinais no Brasil.....	13
1.3 Gênero <i>Pfaffia</i>	13
1.4 Identificação da espécie.....	15
1.5 Identificação de espécies por código de barras de DNA.....	17
1.6 Genes <i>matK</i> e <i>rbcL</i>	20
1.6.1 Gene maturase K.....	20
1.6.2 Gene <i>ribulose-bisphosphate carboxylase</i>	21
1.7 <i>Barcode of life data systems</i>	21
2 OBJETIVOS.....	24
2.1 Objetivo geral.....	24
2.2 Objetivos específicos.....	24
3 METODOLOGIA.....	25
3.1 Coleta de amostras referências.....	25
3.2 Coleta de amostras comerciais.....	25
3.3 Extração de DNA e quantificação.....	25
3.4 Amplificação por PCR.....	26
3.5 Confirmação da amplificação em gel de agarose.....	27
3.6 Purificação do produto de PCR e Sequenciamento.....	28
3.7 Análise dos Resultados.....	29
3.8 Depósito das amostras referências no BOLD.....	30
4 RESULTADOS E DISCUSSÃO.....	31
4.1 Extração de DNA e amplificação por PCR.....	31
4.1.1 Extração de DNA.....	31

4.1.2 Amplificação das amostras referências.....	31
4.1.3 Amplificação das amostras comerciais.....	32
4.2 Sequenciamento e identificação das espécies.....	33
4.2.1 Amostras referências.....	33
4.2.2 Amostras Comerciais.....	35
4.2.3 Utilização dos genes matK e rbcL como código de barras De DNA.....	39
4.3 Importância do BOLD.....	41
5 CONCLUSÃO.....	43
6 REFERÊNCIAS.....	44
ANEXO 1.....	55
ANEXO 2.....	58

1.0 INTRODUÇÃO

1.1 Uso popular de plantas medicinais

O homem, desde sua origem, busca meios de melhorar sua qualidade de vida e conseqüentemente ter uma maior perspectiva da mesma. Um desses meios é a sua interação com a natureza, visando a utilização de recursos naturais para se beneficiar. Uma relação evidente entre seres humanos e plantas, é o uso de diversos recursos vegetais para inúmeras funções, como nos casos da alimentação e das finalidades medicinais, como as plantas medicinais (BALICK & COX, 1997).

A Organização Mundial da Saúde (OMS) define planta medicinal como sendo todo e qualquer vegetal que possui, em um ou mais órgãos, substâncias que podem ser utilizadas com fins terapêuticos ou que sejam precursores de fármacos semissintéticos (WHO, 1998). Já um fitoterápico é o medicamento obtido empregando-se exclusivamente matérias-primas ativas vegetais. É caracterizado pelo conhecimento da eficácia e dos riscos de seu uso, assim como pela reprodutibilidade e constância de sua qualidade (BRASIL, 2004). A diferença entre planta medicinal e fitoterápico reside na elaboração da planta para uma formulação específica, o que caracteriza um fitoterápico (WHO, 1998).

O emprego de plantas medicinais para na saúde tem ocorrido no decorrer do tempo desde as formas mais simples de tratamento local até as formas mais sofisticadas de fabricação industrial de medicamentos (HAMILTON, 2004; LORENZI & MATOS, 2008).

As plantas são uma das mais importantes fontes de novas substâncias bioativas para o desenvolvimento de produtos farmacêuticos e cosméticos. Estima-se que 25% dos medicamentos modernos existentes hoje são derivados de produtos naturais e, no caso dos antitumorais, esta porcentagem chega a 60% (LI & VEDERAS; SAHOO et al., 2009). Segundo a OMS, entre 65% e 80% das populações dos países em desenvolvimento usam atualmente plantas medicinais

como medicamentos e cerca de 85% da medicina tradicional envolve o uso de extratos de plantas, isso significa que 3,5 a 4,0 bilhões de pessoas dependem de plantas como fontes de drogas (FARNSWORTH 1997; WHO, 2011).

O uso de plantas medicinais, assim como a fitoterapia, se encontra em expansão em todo o mundo e constituem um mercado bastante promissor (CALIXTO, 2000; JORDAN et al., 2010; KENNEDY, 2005). Esse setor movimenta aproximadamente US\$160 milhões por ano no Brasil, e mais de US\$21,7 bilhões no mundo (CARVALHO et al., 2011) e passou por uma forte ascensão na última década (FEBRAFARMA, 2009). Com isso, este setor vem atraindo cada vez mais o interesse dos países desenvolvidos, chegando os Estados Unidos a movimentar cerca de US\$ 6,5 bilhões em 2000 e a Europa US\$ 8,5 bilhões, sendo a Alemanha o maior mercado mundial (CARVALHO et al., 2011).

O desenvolvimento de novos produtos a partir de fontes naturais é encorajado visto que se estima que dentre as 300 mil de espécies de plantas existentes no mundo, apenas 15% foram avaliadas para determinar o seu potencial farmacológico (DE LUCA et al., 2012). Além disso, a indústria farmacêutica também tem interesse nas plantas medicinais como fonte de modelo para fármacos (SHU, 1998; HARVEY, 2000) e como matéria prima para desenvolvimento de fitoterápicos (SCHENKEL et al., 2001).

O grande desenvolvimento desse setor de plantas medicinais pode ser atribuído a múltiplos fatores, tais como efeitos adversos de fármacos sintéticos e a preferência dos consumidores por tratamentos considerados naturais. Além disso, o desenvolvimento de novas formas de preparações e administrações de produtos fitoterápicos, a comprovação científica das propriedades farmacológicas de espécies vegetais e seus derivados, um melhor conhecimento químico, farmacológico e clínico, além do menor custo se comparado com os fármacos sintéticos, também contribuem para essa expansão (VIEIRA, 2001; CAÑIGUERAL et al., 2003; WACHTEL-GALOR & BENZIE, 2011).

Devido à utilização generalizada de plantas medicinais, a OMS publicou monografias relativas a plantas medicinais (WHO, 1999-2010). Esses volumes contêm uma lista de espécies com benefícios medicinais reconhecidos e os meios de usá-los corretamente (WHO, 1999; 2003; 2007; 2009; 2010).

1.2 Uso de plantas medicinais no Brasil

O Brasil possui a maior diversidade vegetal do mundo, com cerca de 60.000 espécies vegetais superiores catalogadas (PRANCE, 1977). Destas, apenas 8% foram estudadas para pesquisas de compostos bioativos e 1.100 espécies foram avaliadas em suas propriedades medicinais (GUERRA et al., 2001).

Para regulamentação do uso de plantas medicinais, além de seguir as recomendações da OMS, o Brasil tem a sua própria agência, a Agência Nacional de Vigilância Sanitária (ANVISA). A ANVISA também tem a sua própria lista de espécies aprovadas para a fabricação de medicamentos à base de plantas (ANVISA, 2008).

No Brasil, em 2004, foi aprovada a Resolução nº 48, que dispõe sobre o registro de medicamentos fitoterápicos e, entre outros pontos, abrange as etapas de controle de qualidade da droga vegetal, do produto acabado e da importação de produtos fitoterápicos (ANVISA, 2004).

Somando-se a isso, o Programa Nacional de Plantas Medicinais e Fitoterápicos, instituído em 2007, visa garantir à população brasileira o acesso seguro e o uso racional de plantas medicinais e fitoterápicos, promovendo o uso sustentável da biodiversidade, o desenvolvimento da cadeia produtiva e da indústria nacional. Com vistas a atingir o objetivo desse programa, dentre as proposições, destaca-se a de promover e reconhecer as práticas populares e tradicionais de uso de plantas medicinais, fitoterápicos e remédios caseiros (MINISTÉRIO DA SAÚDE, 2006).

1.3 Gênero *Pfaffia*

As plantas do gênero *Pfaffia* pertencem à família Amaranthaceae e são conhecidas como fafia, corango-sempre-viva ou ginseng brasileiro (NICOLOSO et al., 2001) e o início de sua utilização como medicamento é desconhecida. Plantas das espécies *Pfaffia* spp. apresentam grande volume de vendas no Brasil devido ao seu uso na medicina popular, sendo as *Pfaffia glomerata* e *Pfaffia paniculata* as mais conhecidas e utilizadas (OLIVEIRA, 1986).

Acredita-se que as pessoas do interior do Brasil tenham aprendido com os índios as propriedades de suas raízes, mas foi apenas há alguns anos que a planta começou a ser denominada como Ginseng-brasileiro e ganhou notória popularidade através da mídia (OLIVEIRA, 1986).

Cerca de 90 espécies de *Pfaffia* são conhecidas na América Central e na América do Sul. No Brasil, são descritas 27 espécies encontradas nos estados da Amazônia, do Mato Grosso e de São Paulo (VIDAL & VIDAL, 1967).

O termo “Ginseng” pode se referir a muitas plantas diferentes, como *Panax ginseng*, *Panax japonicus*, *Panax quinquefolium*, *Pfaffia glomerata*, *Pfaffia paniculata*, entre outros, sendo o *Panax ginseng* (coreano) o mais conhecido e estudado. Apesar de amplos estudos com o gênero *Panax*, ainda se faz necessário estudos acerca do gênero *Pfaffia*, tendo em vista sua ampla aplicação como planta medicinal e as diferenças entre esses gêneros (HOSTETTMANN & MARSTON, 1995; OLIVEIRA, 1986).

O Ginseng brasileiro é indicado como tônico e revigorante geral, no tratamento de fadiga física, esgotamento mental, falta de memória, como auxiliar no tratamento de distúrbios circulatórios, como antidiabético, anti-inflamatório, imunoestimulante, dentre outros (OLIVEIRA, 1986).

No entanto, a indicação do Ginseng brasileiro pode diferenciar de acordo com a espécie utilizada. A espécie *Pfaffia glomerata* acelera a cicatrização de úlceras

digestivas (FREITAS et al., 2008) e de feridas na pele (SILVA et al., 2010). Ainda, foi demonstrado que essa planta tem efeito depressor do sistema nervoso (FENNER et al., 2008). Sobre a *Pfaffia paniculata*, foi verificado a presença de atividade antiproliferativa *in vitro* na hepatocarcinogênese (DA SILVA et al., 2010), no câncer de mama (NAGAMINE et al., 2009), além de afetar a atividade de macrófagos (PALERMO et al., 2006) e atuar como afrodisíaco (AELETTI et al., 1999).

Diferenças na composição química entre as espécies podem gerar perfis farmacológicos e toxicológicos diferenciados. Diante disso, é importante que seja feita uma correta identificação da espécie em uso, pois testes biológicos apontam atividades distintas entre as mesmas (TAKEMOTO et al., 1983; SHIOBARA et al., 1992).

A diferenciação das espécies de *Pfaffia* spp. com base apenas em caracteres morfológicos, que no Brasil são mais de 27, não é tarefa simples. A identificação da espécie pode ser laboriosa, pois as plantas medicinais comercializadas encontram-se em sua maioria na forma de pequenos fragmentos ou pulverizadas, o que impossibilita a identificação morfológica. Por esta razão, são encontradas frequentemente no comércio substituição entre as espécies. Este fato constitui um risco à segurança de uso pela população (PALHARES et al., 2014; BRUNI et al., 2010; WHO, 2004).

1.4 Identificação da espécie

Para que ocorra o uso seguro e eficaz da planta medicinal faz-se necessário a correta identificação da espécie. Para garantir a qualidade de medicamentos com plantas medicinais, critérios morfológicos e químicos foram estabelecidos nas Farmacopéias e devem ser seguidos:

- 1) identificação correta das espécies de plantas;
- 2) confirmação da presença e da concentração mínima dos ingredientes ativos;
- 3) análise da pureza.

No item 1 avalia-se se a droga vegetal em análise corresponde à espécie descrita. A segunda etapa é para verificar a presença e teor dos princípios ativos, necessários para as atividades farmacológicas da espécie. O terceiro item consiste na avaliação dos caracteres de pureza a partir da verificação da presença de elementos estranhos, além de verificar a ocorrência de outras adulterações. (BRANDAO et al., 2013).

Existem várias formas de identificação de espécie da planta, sendo que a principal é feita por taxonomistas altamente treinados por meio da observação de características anatômicas e morfológicas do espécime (ANVISA, 2014), que pode ser dificultosa ou impossibilitada devido à elevada variação fenotípica, a comercialização da planta processada em pequenos fragmentos ou pulverizadas e a falta de profissionais treinados, visto que a identificação macroscópica e microscópica da planta requer considerável experiência, principalmente para plantas estreitamente similares (CHEN et al., 2010; DE BOER et al., 2014; KOOL, 2012; NEWMMASTER et al., 2009).

Quando a análise taxonômica não é possível ou para confirmação da mesma, uma alternativa é o teste químico, usando métodos cromatográficos para verificar a presença de substâncias específicas (marcadores químicos) em comparação com um perfil químico do produto na literatura ou encontrados em amostras padrão permitindo a distinção entre as espécies (LI, 2008; TECHEN et al., 2004; WHO, 1998).

Todavia, tais substâncias químicas podem apresentar alguns impasses. Por exemplo, encontrar marcadores químicos que são únicos para as espécies em questão, até porque diferentes espécies podem produzir o mesmo marcador. Além disso, dentro de uma mesma espécie pode ocorrer variabilidade na composição química, dependendo de alguns fatores, como estação do ano, fase de crescimento

em que a planta se encontra, colheita, armazenamento e processamento (BENNETT, R. N. & WALLSGROVE, 1994).

Além disso, a segurança dos produtos à base de plantas pode ser comprometida por erros de identificação, contaminação deliberada e adulteração acidental (ERNST, 2004; BREEMEN et al., 2008), o que pode levar a graves efeitos colaterais, devido à presença de outros compostos que não a planta em questão. Isto cria a necessidade de autenticação de espécies incluída nesses produtos medicinais e novas formas de identificação para assim aumentar as possibilidades de identificação (GILBERT, 2011).

As limitações do sistema de identificação morfológica e a insuficiência de taxonomistas especializados nos diferentes grupos de organismos impulsionaram o desenvolvimento de novos métodos para a identificação de espécies como, por exemplo, a identificação por análise do DNA (AZEREDO, 2005).

1.5 Identificação de espécies por código de barras de DNA

Atualmente, tem sido desenvolvido técnicas alternativas na identificação de espécies. Uma delas, é o uso de um sistema proposto no “Projeto do Código de Barras da Vida” (*The Barcode of Life Project*), iniciado em 2003 no Canadá, e que consiste em um sistema único e universal com o objetivo de identificar todos os organismos em nível específico, baseado em uma abordagem molecular padronizada, utilizando-se sequências de regiões específicas do DNA (AZEREDO, 2005; HEBERT et al., 2003).

Esse projeto foi promovido a partir de 2004 através do *Consortium for the Barcoding of Life* (CBOL) (KOSMANN, 2009), devido ao rápido progresso desta técnica. O

CBOL objetiva a geração de código de barras genéticos para todas as espécies de seres vivos do planeta (RATNASINGHAM & HEBERT, 2007).

A partir deste projeto foi criado um banco de dados mundial (*Barcode of Life Data Systems* - BOLD - <http://www.boldsystems.org>), que consiste em uma plataforma de bioinformática que contém sequências das mais diversas espécies, tanto animais, vegetais, fungos e outros organismos (RATNASINGHAM & HEBERT, 2007). O BOLD, além de conter as sequências, garante a qualidade das mesmas. Por meio deste banco de dados, será possível identificar qualquer espécie apenas comparando a sequência de DNA de interesse com as sequências depositadas no mesmo (Figura 1) (AZEREDO, 2005).

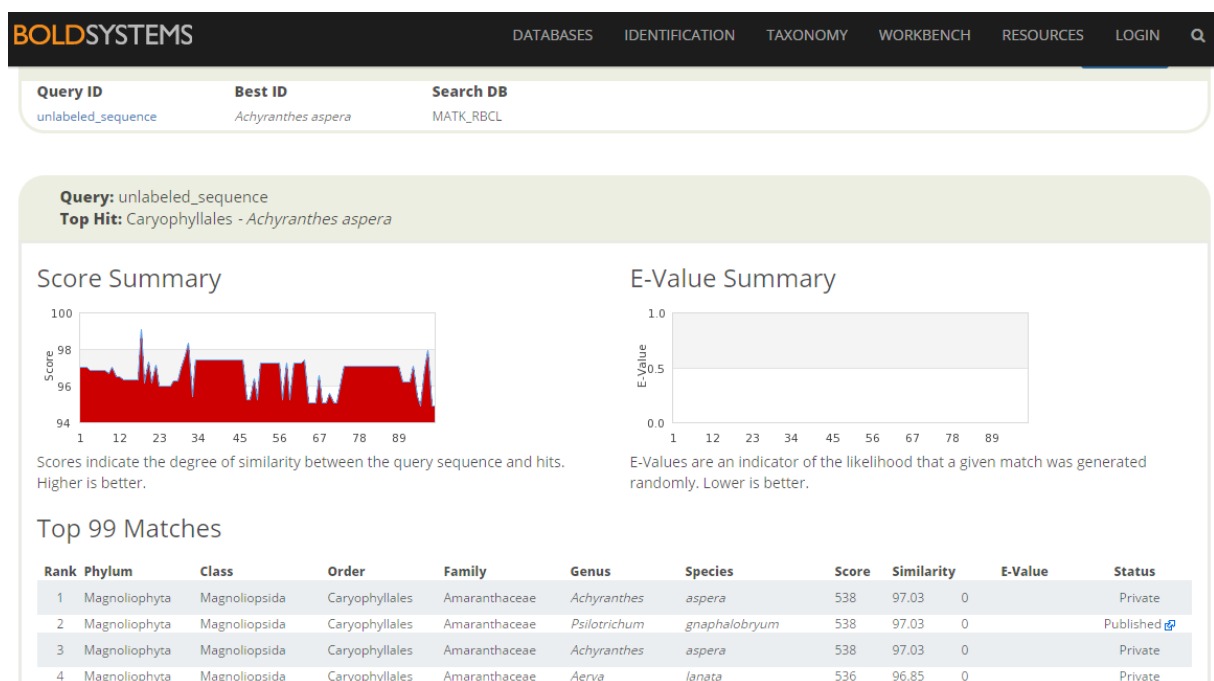


Figura 1 – Resultado de uma busca do gene *rbcL* da planta *Pfaffia glomerata* no BOLD (<http://v4.boldsystems.org/index.php>). Fonte: Autor

Além de fornecer uma ferramenta de identificação rápida, com apenas uma pequena quantidade de tecido de qualquer fase de desenvolvimento de uma planta ou de um animal, o projeto de códigos de barra da vida também contribui com a descoberta de novas espécies (HEBERTH et al., 2004a; HEBERTH et al., 2004b).

Para que seja possível a realização dessa técnica de DNA *Barcoding*, deve-se encontrar uma região ideal no DNA e que atenda à algumas especificações. O fragmento deve ter sequência com comprimento curto para facilitar a amplificação e o sequenciamento, ser suficientemente variável para separar espécies muito parecidas e apresentar sítios conservados para o desenvolvimento de *primers* universais (FORD et al., 2009).

A proposta é de que qualquer pessoa, em qualquer lugar do mundo, e a qualquer hora, seja capaz de identificar de forma rápida e eficiente a espécie de um determinado espécime, independente da condição de preservação do ser em questão (STOECKLE et al. 2005).

Para a identificação de animais por código de barras de DNA, é empregada a região de 648 pares de base do gene COI do DNA mitocondrial (Citocromo C Oxidase subunidade I, tendo como referência o genoma mitocondrial de *Mus musculus Linnaeus*, 1758). Esse marcador foi escolhido por essa sequência funcionar como um código de barras exclusivo para cada espécie, tendo como premissa que a variação interespecífica é bem mais elevada que a intraespecífica (de até 3%) (HEBERT et al., 2003).

Em plantas, entretanto, a baixa taxa de substituição do DNA mitocondrial levou a procura por regiões de código de barras alternativas, como as regiões de plastídios (CBOL, 2009; FAZEKAS et al., 2008; HAJIBABAEI et al., 2008; KRESS et al., 2005). Na sequência, vários estudos baseados em diferentes loci dos cloroplastos e genomas nucleares levaram à conclusão de que nenhum locus poderia, sozinho, ser adotado como código de barras em plantas e logo percebeu-se que os códigos de barras de multilocus são requisito para sua identificação (PENNISI, 2007; LEDFORD, 2008).

Posteriormente, um certo número de loci foram testados quanto à sua adequação como códigos de barras de plantas e muitas combinações de multilocus foram sugeridas. Das investigações iniciais das regiões dos plastídios, sete principais candidatos surgiram (PENNISI, 2007; LEDFORD, 2008). Quatro são porções de genes de codificação (matK, rbcL, rpoB e RpoC1) e 3 são espaçadores não codificadores (atpF-atpH, trnH-psbA, E psbK-psbI). Diferentes grupos de pesquisa propuseram várias combinações destes loci como seus códigos de barras preferidos

da planta, mas nenhum consenso surgiu (LAHAYE, 2008; ERICKSON et al., 2008; HOLLINGSWORTH, 2009; SEBERG & PETERSEN, 2009).

Dentre as várias combinações de locus, algumas são sugeridas, dentre elas: RpoC1, rpoB, matK (CHASE et al., 2005); rpoC1, matK, trnH-psbA; rbcL e trnH-psbA (KRESS & ERICKSON, 2007). O CBOL, 2009, recomendou a combinação de dois genes, o matK (maturase K gene) e o rbcL (ribulose-bisphosphate carboxylase gene) como códigos de barras de plantas.

Os maiores impasses encontrados no emprego do DNA *Barcoding* são: resolução de espécies que divergiram recentemente, grupos com baixa diversidade gênica em suas sequências detecção de híbridos e os pseudogenes nucleares (Numts) (STOECKLE et al. 2005).

O banco de dados dispõe, atualmente, de mais de quatro milhões de amostras depositadas englobando cerca de 265 mil espécies de plantas (dados de novembro de 2016). Além disso, uma rede de cientistas brasileiros integra o *Brazilian Barcode of Life* (BrBOL - <http://www.brbol.org>) cuja proposta é gerar dados de espécies brasileiras para o banco de dados, vinculado ao BOLD.

1.6 Genes matK e rbcL

1.6.1 Gene maturase K

O gene *maturase K* (matK) está localizada dentro de um íntron classe II, entre os exons 5' e 3' do gene que codifica o trnH. Utiliza-se para o código de barras de DNA o tamanho de 841 pb, localizado no centro do gene matK.

Este gene apresenta uma das mais rápidas taxas de substituição dentre as regiões codificantes dos cloroplastos. Assim sendo, a taxa de evolução do matK é adequada para resolver relações intergênero, bem como as relações interespecies em muitas plantas angiospermas (CBOL, 2009; HILU e LIANG, 1997; HOLLINGSWORTH et al., 2011).

1.6.2 Gene *ribulose-bisphosphate carboxylase*

O gene *ribulose-bisphosphate carboxylase* (rbcL) é codificante de uma proteína altamente conservada, a ribulose-1,5-bisfosfato, conhecida como Rubisco. Possui um tamanho de aproximadamente 1400 pb (FORD et al., 2009), porém na técnica de DNA *Barcode*, utiliza-se um fragmento menor, com aproximadamente 550 pb.

Sua escolha como *barcode* foi adotada por ser um gene comum entre as plantas e pela sua utilização histórica em estudos filogenéticos (STOECKLE et al., 2011). Apresenta uma taxa de evolução lenta e bastante conservada, apresentando boa resolução em níveis taxonômicos mais elevados, como família (CHASE et al., 2007).

Uma alta universalidade, mas com pouca resolução entre as espécies é fornecida pelo rbcL, enquanto matK oferece alta resolução, mas menos universalidade. A combinação destes dois genes pode ajudar a alcançar maior discriminação entre espécies. No BOLD pode-se inserir a sequência a ser analisada para comparação com outras espécies já depositadas, tanto para matK quanto para rbcL (SINGH et al., 2012; CAMERON et al., 1999).

1.7 Barcode of life data systems



Figura 2 – Site do banco de dados BOLD (<http://v4.boldsystems.org/index.php>).

Quando se utiliza a técnica de DNA *barcoding* é importante que esse registro seja feito no banco de dados, como o BOLD (Figura 2), para que a espécie identificada sirva de “molde” para identificações posteriores. Para que o espécime seja incluso no BOLD, algumas especificações devem ser cumpridas para que o depósito seja de total confiabilidade (RATNASINGHAM & HEBERT, 2007):

- 1) nome da espécie;
- 2) dados do material;
- 3) testemunho (número e instituição de depósito);

- 4) dados de coleta (nome do coletor, data e local da coleta com coordenadas geográficas);
- 5) nome do identificador do espécime;
- 6) sequência do DNA com, no mínimo, 500 pb;
- 7) *primers* utilizados;
- 8) eletroferogramas.

Portanto, é essencial que os bancos de dados estejam atualizados para possibilitar sua utilização da melhor maneira possível. Quanto maior o número de espécies depositadas maiores as possibilidades de identificação (HEBERT & GREGORY, 2005). O objetivo principal do DNA *Barcoding* é disponibilizar uma metodologia rápida, precisa para identificação das espécies, utilizando regiões do DNA padronizadas (HEBERT, 2003).

2 OBJETIVOS

2.1 Objetivo geral

Identificar, em nível de espécie, amostras de Ginseng-brasileiro comercializadas no mercado brasileiro, utilizando, para isso, o código de barras de DNA (*DNA barcode*).

2.2 Objetivos específicos

- Caracterizar amostras referência morfológicamente identificadas das espécies *Pfaffia glomerata* e *Pfaffia paniculata*;
- Caracterizar amostras comerciais vendidas no mercado brasileiro como Ginseng-brasileiro;
- Extrair o DNA, amplificar e sequenciar os genes *matK* e *rbcL* de todas as amostras;
- Depositar no banco de dados BOLD, sequências dos genes *matK* e *rbcL* para as amostras referência das espécies *P. glomerata* e *P. paniculata*;
- Verificar qual o potencial da utilização do código de barras de DNA na identificação das espécies *Pfaffia paniculata* e *Pfaffia glomerata*.

3 METODOLOGIA

3.1 Coleta de amostras referências

Foram obtidas 30 amostras referências, da Universidade Estadual de Campinas (Unicamp) - SP (n=13) e Universidade Federal de Mato Grosso (UFMT) - MT (n=17). Dessas amostras, 6 foram *P. paniculata* (SP) e o restante *P. glomerata* (SP e MT). As amostras se encontravam em forma de raiz ou folha seca e foram corretamente identificadas morfologicamente.

3.2 Aquisição de amostras comerciais

Amostras vendidas como ginseng brasileiro (n = 60) de diversos locais do país foram coletadas por meio de compra física ou internet (Anexo 1).

3.3 Extração de DNA e quantificação

Para extração de DNA utilizou-se o kit Nucleo Spin® Plant II (Macherey-Nagel), seguindo o seu protocolo, utilizando 20 mg de amostra. Para isso, foi necessário macerar algumas amostras com auxílio de nitrogênio líquido, visto que as mesmas se encontravam em forma de raiz ou pequenos pedaços, sendo que era necessária na forma de pó fino para extração de DNA.

Posteriormente à extração, as amostras foram submetidas à quantificação em NanoDrop 2000c (Thermo Scientific).

3.4 Amplificação por PCR

As amplificações por PCR foram feitas em duas regiões do cloroplasto, *matK* e *rbcl*. A tabela 1 apresenta os *primers* utilizados.

Tabela 1 – *Primers* utilizados neste estudo.

Gene	Sentido	Nome	Sequencia 5' – 3'	Temperatura de anelamento
<i>matK</i>	Forward	MatK-1RKIM-f	ACCCAGTCCATCTGGAAATCTT GGTTC	52°C
<i>matK</i>	Reverse	MatK-3FKIM-r	CGTACAGTACTTTTGTGTTTACG AG	52°C
<i>rbcl</i>	Forward	<i>rbcl</i> a-F	ATGTCACCACAAACAGAGACTA AAGC	55°C
<i>rbcl</i>	Reverse	<i>rbcl</i> a-R	GTAAAATCAAGTCCACCRCG	55°C

A reação foi padronizada utilizando os seguintes reagentes e quantidades (Tabela 2):

Tabela 2 – Reagentes e quantidade utilizados na reação de PCR

Reagentes	Concentração
Buffer 10X Invitrogen 200mM Tris-HCl (pH 8,4), 500mM KCl;	0,25 mM
MgCl ₂ 50 mM (Invitrogen)	0,03 mM
<i>Primer</i> foward 1,25 picomolar (pM)	1,25 pM
<i>Primer</i> reverse 1,25 pM	1,25 pM
10 mM dNTP mixture (Promega)	6,25 x 10 ⁻⁴ mM
Taq DNA Polimerase (Invotrogen)	0,3U

A quantidade de DNA utilizada na reação de PCR variou entre 30-50 nanogramas (ng). (ex. O volume total da reação foi de 12,25 µl, sendo que a quantidade dos reagentes citados na Tabela 1 foi acrescida do DNA e finalizada com água ultrapura para integralizar o volume).

Em cada reação de PCR realizada foram utilizados controles positivos (amostra referência) e negativos (água ultrapura), para assim garantir a qualidade da reação de PCR.

Utilizou-se o termociclador (Applied Biosystems Veriti Thermal Cycler) conforme o ciclo abaixo:

matK:

94°C - 1 min

94°C - 30 seg / 52°C - 20 seg / 72°C - 1 min – 35 ciclos

72°C - 5 min

4°C - ∞

rbcL

94° - 4 min

94°C - 30 seg / 55°C - 30 seg / 72°C - 1 min – 35 ciclos

72°C - 10 min

4°C - ∞

Em algumas reações de PCR, na segunda tentativa de amplificação, foi adicionado DMSO (Dimetilsulfóxido) 5%, como alternativa para aumentar as taxas de amplificação das reações.

3.5 Confirmação da amplificação em gel de agarose

Para confirmação da amplificação do fragmento foi utilizada a técnica de eletroforese em gel de agarose 1,5%, por aproximadamente 50-60 minutos a 100 volts, utilizando 4 µl do produto de PCR. Os fragmentos foram comparados a um Ladder RTU

(Ready-to-use KASVI). As avaliações foram feitas visualmente, por meio da coloração com Brometo de Etídeo, utilizando luz Ultravioleta em fotodocumentador.

3.6 Purificação do produto de PCR e Sequenciamento

Primeiramente os produtos de PCR foram purificados com enzimas Exonuclease I (Exo I) e Shrimp Alkaline Phosphatase (SAP) para retirada de reagentes excedentes da reação de PCR, como *primers* e nucleotídeos não incorporados. Para cada reação de PCR foi utilizado 1 µl de ExoSap diluída (1:4 ExoSap e água ultrapura, respectivamente). O aparelho utilizado foi o termociclador (Applied Biosystems Veriti Thermal Cycler) com o ciclo:

37°C - 30 minutos

80°C - 15 minutos

4°C - ∞

Posteriormente foram submetidos a reação de sequenciamento conforme o protocolo do fabricante do Big Dye Terminator v. 3.1 (*Applied Biosystems*) (Tabela 3):

Tabela 3 – Reagentes e quantidades utilizados para reação de sequenciamento.

Reagentes	Concentração
Save Money (Tris-HCl 1M e MgCl ₂ 50mM)	0,035 M
Água ultrapura autoclavada qsq. 10 µL	2,0 µl
<i>Primer</i> 1,6 pM	3,2 pM
Big Dye Terminator v. 3.1	0,5 µl

Foram adicionados 2 µl do produto de PCR purificado.

O mesmo termociclador citado acima foi utilizado para tal reação, seguindo o ciclo abaixo:

96°C - 1 min

96°C - 15 seg / 50°C - 15 seg / 60°C - 4 min – 35 ciclos

8°C - ∞

Após ciclagem, os produtos do sequenciamento foram precipitados utilizando o método EDTA/Acetato de sódio/Etanol descrito a seguir:

Pipetou-se 1 µl de EDTA 125mM em cada amostra e vortexou gentilmente. Adicionou 1 µl de acetato de sódio 3M e vortexou novamente. Pipetou-se 32 µl de álcool etílico 95%, vortexou e manteve-se à temperatura ambiente por 15 minutos e posteriormente centrifugado a 2500 RCF por 30 minutos. Imediatamente após a centrifugação, foi removido o sobrenadante e adicionou-se 35 µl de álcool etílico 70% e vortexou-se. Novamente foi centrifugado a 2500 RCF por 15 minutos. Logo após, removido o sobrenadante. Os tubos foram levados ao termociclador por 10 minutos a 40° C para completa evaporação do álcool.

Por fim, as reações foram analisadas em sequenciador automático ABI 3500 (Applied Biosystems) do Laboratório do Núcleo de Genética Aplicada à Conservação da Biodiversidade, da Universidade Federal do Espírito Santo, conforme recomendações do fabricante.

3.7 Análise dos Resultados

Os eletroferogramas foram analisados utilizando o programa BioEdit (Hall, 1999). As amostras em análise foram alinhadas e confrontadas com as amostras referência, para identificação das espécies.

Os resultados também foram comparados as informações disponibilizadas nos rótulos das amostras e também foram confrontados com o BOLD.

Com os alinhamentos das amostras referências, gerados no BioEdit, foram construídas árvores filogenéticas através do programa MEGA7, método Neighbor-Joining (NJ).

3.8 Depósito das amostras referências no BOLD

Todas as amostras referências foram depositadas no BOLD. Primeiramente, foram adicionadas as informações a respeito da espécie e posteriormente sobre o gene em questão. Posteriormente a sequência obtida com o sequenciamento foi adicionada.

4 RESULTADOS E DISCUSSÃO

4.1 Extração de DNA e amplificação por PCR

4.1.1 Extração de DNA

Todas as amostras tiveram os DNAs extraídos utilizando um kit próprio para plantas e foram quantificadas em Nanodrop previamente à amplificação. As amostras referência resultaram em uma média de 31 ng/uL de DNA e pureza (relação 260/280) média de 2,2. As amostras comerciais resultaram em uma média de 12 ng/uL de DNA e pureza média de 1,9 e foram submetidas à amplificação.

4.1.2 Amplificação das amostras referências

As taxas de amplificações para as amostras referências foram inicialmente 80% (24 amostras) e 83,33%(25 amostras), para os genes *matK* e *rbcl*, respectivamente. Posteriormente, adicionou-se DMSO à reação de PCR, as taxas aumentaram para 90% (27 amostras), tanto para o gene *matK* quanto para o *rbcl*.

Dantas et al. (2002), Takiuchi et al. (2003), Ralser (2006), também utilizaram DMSO para melhorar a eficiência das reações de PCR, e todos obtiveram êxito, aumentando consideravelmente o rendimento e qualidade da reação.

Vários compostos podem ser utilizados para aumentar o rendimento, especificidade e consistência da reação de PCR, sendo que o DMSO é um dos mais utilizados. Ele age facilitando a separação das cadeias de DNA, interferindo na ligação das bases (FRACKMAN et al., 1998, JENSEN et al., 2010). O uso de DMSO promove, portanto, a desestabilização da fita dupla de DNA, facilitando o anelamento dos *primers* (MASOUD et al., 1992).

Para as amostras referências que não amplificaram foram repetidas as reações de PCR, porém sem sucesso.

4.1.3 Amplificação das amostras comercializadas

Todas as 60 amostras comercializadas disponíveis para análise foram submetidas à amplificação por PCR, utilizando os *primers* específicos para os genes *matK* e *rbcL*. As taxas de amplificações iniciais foram de 86,67% (52 amostras) para o gene *matK* e 76,67% (46 amostras) para o gene *rbcL*. Após a adição de DMSO na reação de PCR as taxas de amplificação aumentaram para 95% (57 amostras) para ambos os genes. Para as amostras que não amplificaram, uma nova tentativa de PCR foi realizada, entretanto, não obtivemos êxito.

Palhares et al. (2015), obtiveram taxa de sucesso de 81,32% e 88,72%, para *matK* e *rbcL*, respectivamente, sendo que seu trabalho englobava diversas espécies. Já Januario (2014), obteve a taxa de amplificação para o gene *rbcL* inferior, 77,77%, para as espécies *Senna Mill. (Fabaceae)* e *Casearia Jacq (Salicaceae)*, quando comparado ao estudo de Kress e Erickson (2007), que relataram taxa de amplificação de 95,8% e 92,70%, para os diferentes gêneros estudados.

Para Palhares et al. (2014), a taxa de sucesso na PCR foi de 65,22% para o *matK* e 100% para o *rbcL*, em estudo sobre *Cinchona*. O grupo chinês BOL de plantas e colaboradores (2011), encontraram um resultado de amplificação de 91,00% para

matK e 94,50% para rbcL, para diversificadas espécies. Portanto a taxa de sucesso encontrada no presente trabalho está em conformidade com o encontrado na literatura.

4.2 Sequenciamento e identificação das espécies

4.2.1 Amostras referências

Todas as amostras amplificadas por PCR (90%, n=27) foram sequenciadas, nas fitas *forward* e *reverse*, tanto para o gene matK, quanto para o rbcL. Destas, 92% geraram eletroferogramas de boa qualidade para o gene rbcL e 90% para o gene matK. Entretanto, nem todas as amostras sequenciadas geraram um sequenciamento de boa qualidade. Para o gene rbcL, 5 amostras geraram eletroferogramas de baixa qualidade e 1 amostra para o gene matK.

As amostras referência de *P. glomeratas* e *P. paniculatas* mostraram-se idênticas, sem nenhuma base nucleotídica diferente entre elas, mesmo sendo obtidas de fontes diferentes. Essa ocorrência pode ser explicada por não haver diferença intraespecífica ou a amostragem não foi significativa (Tabela 4). Em anexo árvore filogenética (Anexo 2).

Tabela 4 – Taxas de variações intraespecífica e interespecífica *P. glomerata* e *P. paniculata* encontradas em amostras referência deste trabalho.

Gene	Varição intraespecífica	Varição interespecífica
matK	0%	1,92%
rbcL	0%	1,90%

As amostras referências foram sequenciadas e alinhadas utilizando o programa BioEdit. O tamanho do fragmento gerado para matK foi de cerca de 1000pb, porém foram desconsiderados o início e o fim da sequência, restando 830 pb para análise. Para o rbcL o tamanho de fragmento gerado foi de 620 pb, desconsiderando o início e o fim, foram analisados 525 pb.

Quando as sequências das duas espécies referência de *Pfaffia* foram comparadas, foram encontradas diferenças nucleotídicas de 16 bases entre a *P. glomerata* e a *P. paniculata* para o gene matK, o que representa em torno de 1,92% de variação interespecífica. Para o rbcL, foram encontradas 10 bases diferentes ao longo da sequência, portanto apresenta cerca de 1,90% de discrepância entre as espécies em estudo (Tabela 4). Saslis-Lagoudakis et al. (2015), encontraram 36 substituições (25 para matK e 11 para trnH-psbA) nas amostras de *Equisetum palustre* e *Equisetum arvense*, sendo que algumas delas foram exclusivas para cada espécie e outras foram compartilhadas com outras espécies.

Além disso, em alguns casos, não é possível a diferenciação entre espécies porque a variação genética dentro do gênero não é suficiente (menor do que 1%), como o relatado por Palhares et al. (2015). No nosso estudo, entretanto, as espécies possuem variações acima deste valor, o que possibilita a identificação da espécie com ambos os genes.

As árvores filogenéticas geradas mostram a distância genética das amostras entre si e relacionadas a outro gênero, o *Panax* (Anexo 2).

4.2.2 Amostras Comerciais

Todas as amostras comerciais amplificadas foram submetidas ao sequenciamento (95%, n=57). Destas, 80% geraram eletroferogramas de boa qualidade para o gene *rbcL* e 75% para o gene *matK*. As amostras comerciais foram classificadas em *P. glomerata*, *P. paniculata* ou não identificadas. Para serem consideradas *P. glomerata* ou *P. paniculata*, as amostras precisavam apresentar 100% de similaridade com a amostra referência, em relação ao gene *matK* e *rbcL*. Foi considerado 100% de similaridade pois as amostras referências não apresentaram variação intraespecífica.

Existem vários testes que podem ser feitos para analisar os resultados de DNA *Barcode* e obter a identificação de espécies. Um dos métodos mais empregados é o “barcoding gap”, em que as espécies são consideradas identificadas se a diferença interespecífica for maior que a diferença intraespecífica (BARRETT & HEBERT, 2005; CBOL, 2009).

As amostras que apresentavam variações em relação à amostra referência foram classificadas como não identificadas. Todas as amostras classificadas como tal apresentaram diferenças de pb de no mínimo 2% com as referências.

Na tabela 5, pode-se observar a denominação de cada amostra comercial em seu rótulo e o resultado do sequenciamento confrontado com as amostras referências.

Tabela 5 – Identificação das amostras comerciais após confronto com as amostras referências.

IDENTIFICAÇÃO DAS AMOSTRAS COMERCIAIS APÓS CONFRONTO COM AMOSTRAS REFERÊNCIAS			
NÚMERO	NOME COMERCIAL NA EMBALAGEM	<i>matK</i>	<i>rbcL</i>
1	<i>Pfaffia</i> (ginseng brasileiro)	<i>P. glomerata</i>	<i>P. glomerata</i>
2	Ginseng (<i>Pfafia paniculata</i>)	X	X
3	Ginseng Brasileiro	X	Não identificada
5	Ginseng	<i>P. paniculata</i>	<i>P. paniculata</i>
6	Ginseng	<i>P. glomerata</i>	<i>P. glomerata</i>
7	Ginseng (<i>Pfafia paniculata</i>)	<i>P. glomerata</i>	<i>P. glomerata</i>
8	Ginseng	<i>P. paniculata</i>	<i>P. paniculata</i>
9	Ginseng	<i>P. glomerata</i>	<i>P. glomerata</i>
10	Ginseng	<i>P. paniculata</i>	<i>P. paniculata</i>
11	Ginseng (<i>Pfafia paniculata</i>)	Não identificada	Não identificada
12	Ginseng	<i>P. glomerata</i>	<i>P. glomerata</i>
13	<i>Panax</i>	<i>P. paniculata</i>	<i>P. paniculata</i>
14	Ginseng	<i>P. paniculata</i>	<i>P. paniculata</i>

Tabela 5 – Identificação das amostras comerciais após confronto com as amostras referências.

IDENTIFICAÇÃO DAS AMOSTRAS COMERCIAIS APÓS CONFRONTO COM AMOSTRAS REFERÊNCIAS			
NÚMERO	NOME COMERCIAL NA EMBALAGEM	matK	rbcl
15	Ginseng	<i>P. paniculata</i>	<i>P. paniculata</i>
16	Ginseng (<i>Pfaffia</i> spp)	<i>P. glomerata</i>	<i>P. glomerata</i>
17	Ginseng	<i>P. glomerata</i>	<i>P. glomerata</i>
18	Ginseng (<i>Pfafia paniculata</i>)	<i>P. glomerata</i>	<i>P. glomerata</i>
19	Ginseng em Pó	<i>P. glomerata</i>	<i>P. glomerata</i>
20	Ginseng Pó Varejo	<i>P. glomerata</i>	<i>P. glomerata</i>
21	Ginseng Chinês (<i>Panax Ginseng</i>)	Não identificada	Não identificada
22	Ginseng	<i>P. glomerata</i>	<i>P. glomerata</i>
23	Ginseng (<i>Pfafia paniculata</i>)	<i>P. paniculata</i>	<i>P. paniculata</i>
24	Ginseng	<i>P. paniculata</i>	<i>P. paniculata</i>
25	Fafia (<i>Pfaffia glomerata</i>)	<i>P. glomerata</i>	<i>P. glomerata</i>
26	Ginseng	<i>P. paniculata</i>	<i>P. paniculata</i>
27	Ginseng	<i>P. paniculata</i>	<i>P. paniculata</i>
28	Ginseng	Não identificada	Não identificada
29	Ginseng	<i>P. glomerata</i>	<i>P. glomerata</i>
30	Ginseng	<i>P. glomerata</i>	<i>P. glomerata</i>
31	Ginseng	<i>P. glomerata</i>	<i>P. glomerata</i>
32	Ginseng Brasileiro (<i>Pfaffia paniculata</i>)	<i>P. glomerata</i>	<i>P. glomerata</i>
33	Ginseng	<i>P. glomerata</i>	<i>P. glomerata</i>
34	Ginseng	<i>P. glomerata</i>	<i>P. glomerata</i>
35	Ginseng	<i>P. glomerata</i>	<i>P. glomerata</i>
36	Ginseng (<i>Pfaffia paniculata</i> Mart.)	<i>P. glomerata</i>	<i>P. glomerata</i>
37	Ginseng	Não identificada	Não identificada
38	Ginseng	<i>P. glomerata</i>	<i>P. glomerata</i>
39	Ginseng (<i>Pfaffia</i>)	<i>P. glomerata</i>	<i>P. glomerata</i>
40	Ginseng	<i>P. glomerata</i>	<i>P. glomerata</i>
41	Ginseng Brasileiro (<i>Pfaffia paniculata</i>)	<i>P. glomerata</i>	<i>P. glomerata</i>
42	Korean Ginseng Tea gold	X	X
43	Ginseng	<i>P. glomerata</i>	<i>P. glomerata</i>
44	Ginseng	<i>P. glomerata</i>	<i>P. glomerata</i>
45	Ginseng	Não identificada	Não identificada
46	Ginseng	<i>P. paniculata</i>	X
49	Ginseng	<i>P. paniculata</i>	<i>P. paniculata</i>
50	Ginseng (<i>Pffafia paniculata</i>)	Não identificada	Não identificada
51	Ginseng em pó	<i>P. paniculata</i>	<i>P. paniculata</i>
52	Ginseng pó	<i>P. paniculata</i>	<i>P. paniculata</i>
53	Ginseng em pó	<i>P. glomerata</i>	<i>P. glomerata</i>
54	Diversos	<i>P. paniculata</i>	<i>P. paniculata</i>
55	Ginseng em pó KG	Não identificada	Não identificada
56	Ginseng em pó	<i>P. paniculata</i>	<i>P. paniculata</i>
57	Ging Seng	Não identificada	Não identificada
58	Ginseng raiz	Não identificada	Não identificada

Tabela 5 – Identificação das amostras comerciais após confronto com as amostras referências.

IDENTIFICAÇÃO DAS AMOSTRAS COMERCIAIS APÓS CONFRONTO COM AMOSTRAS REFERÊNCIAS			
NÚMERO	NOME COMERCIAL NA EMBALAGEM	matK	rbcL
59	Ginseng em po 12	Não identificada	Não identificada
60	Ginseng pó granel	Não identificada	Não identificada
61	Ginseng pó	Não identificada	Não identificada
62	Ginseng Granel	<i>P. glomerata</i>	<i>P. glomerata</i>
65	Ginseng	<i>P. paniculata</i>	<i>P. paniculata</i>

Legenda: X – amostras que não amplificaram.

A utilização dos genes matK e rbcL possibilitou a identificação de grande parte das amostras (67,24%) e a confirmação das amostras que não são nenhuma das duas espécies em questão (32,76%). Diferentemente do resultado apresentado por Zhang e colaboradores (2015), em que a combinação rbcL e matK gerou uma taxa de identificação de diferentes espécies de plantas medicinais de apenas 40,40%. O código de barras rbcL e matK discriminou apenas 49,7% das amostras no trabalho de China Plant BOL Group e colaboradores (2011).

Todas as amostras do estudo foram confrontadas no banco de dados BOLD e no NCBI (apenas para conferência) e não foram identificadas (igualdade de 100%) com nenhuma das espécies já depositadas. Entretanto, uma similaridade de 100% não é exigida para a identificação de todas as espécies de plantas depositadas no BOLD, tendo em vista que podem existir variações maiores que 1% em outros gêneros.

Abaixo é apresentada a tabela 6, expondo a similaridade de algumas amostras não identificadas neste trabalho, mas que apresentaram similaridade acima de 98% com outras espécies depositadas no BOLD.

Tabela 6 – Similaridade das amostras não identificadas com espécies depositadas no BOLD.

Amostra	Identificação no BOLD	Similaridade	Gene
03	<i>Musa acuminata</i> (bananeira)	98,67%	rbcL
21	<i>Withania somnifera</i> (ginseng indiano)	98,71% / 99,49%	rbcL e matK
28	<i>Tribulus terrestris</i>	98,32%	rbcL
55	<i>Iresine calea</i>	98,51%	rbcL
59	<i>Phaseolus vulgaris</i> (feijão)	98,85%	rbcL

Das 58 amostras avaliadas neste estudo, 67,24% foram identificadas, sendo 61,54% *P. glomerata* e 38,46% *P. paniculata*. O restante, 32,76% (19 amostras) não foram identificadas conforme o rótulo, caracterizando, portanto adulteração. Das desconformes com o rótulo, 22,41% (13 amostras) não foram identificadas e 10,34% (6 amostras) apresentaram substituição de espécies, como pode ser visto no gráfico 1.

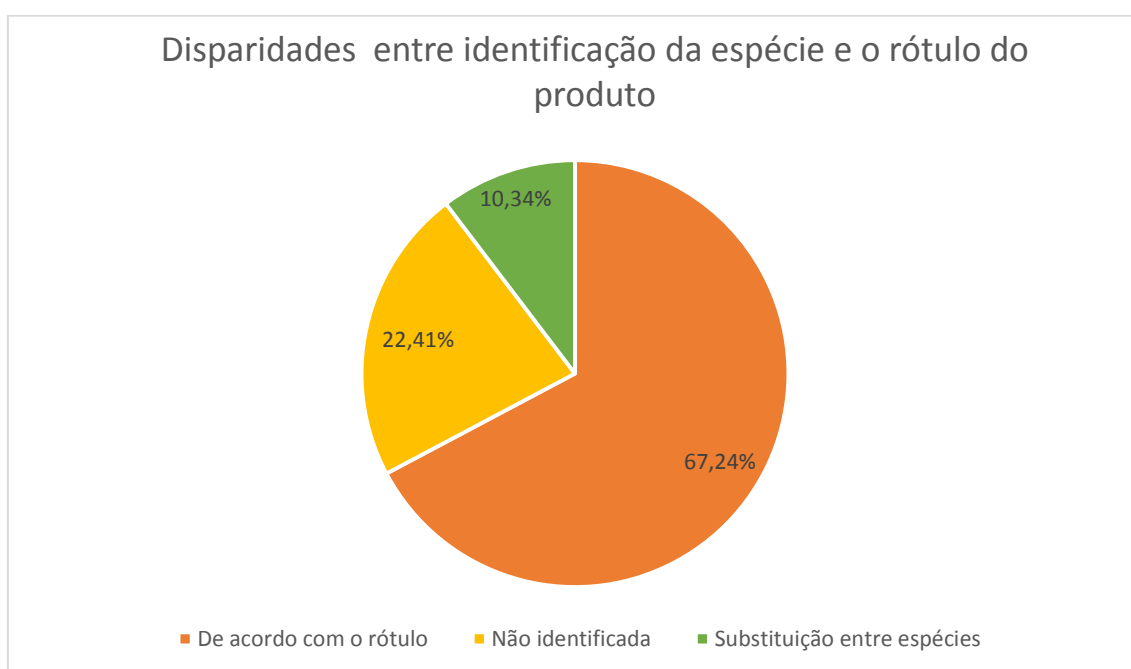


Gráfico 1 – Disparidades encontradas entre a identificação genética da espécie e o rótulo do produto obtidas neste estudo.

Esse resultado mostra-se menor do que o apresentado por Brandão e colaboradores (2013), em que constataram que 50% das 252 amostras de 40 espécies nativas da flora brasileira e que são amplamente comercializadas em mercado nacional, eram na realidade espécies identificadas incorretamente. Han e colaboradores (2016), constataram uma taxa de 4,2% de adulteração em plantas medicinais dos mercados chineses, em que as amostras identificadas não estavam em conformidade com o nome no rótulo.

A amostra 21, que no rótulo a identificava como *Panax*, não foi identificada como nenhuma das duas espécies do estudo. Confrontando com o banco de dados BOLD

e NCBI, também não foi identificada com nenhuma outra espécie em 100%, inclusive nenhuma do gênero *Panax*. Porém obteve semelhança de 98,71% para *rbcL* e 99,49% para *matK* com o *Withania somnifera* (ginseng indiano).

A amostra 13, também rotulada como *Panax*, na realidade tratava-se de *P. paniculata*. Palhares e colaboradores (2015) também encontram substituição do gênero *Panax* por *Pfaffia* em 100% das amostras consideradas anteriormente como *Panax*.

Em uma análise dos eletroferogramas de 6 amostras, pode-se observar a provável mistura de espécies, visto que apresentam picos duplos em diversos locais, tanto para o gene *rbcL*, como para o *matK*, como pode ser observado na figura 3.

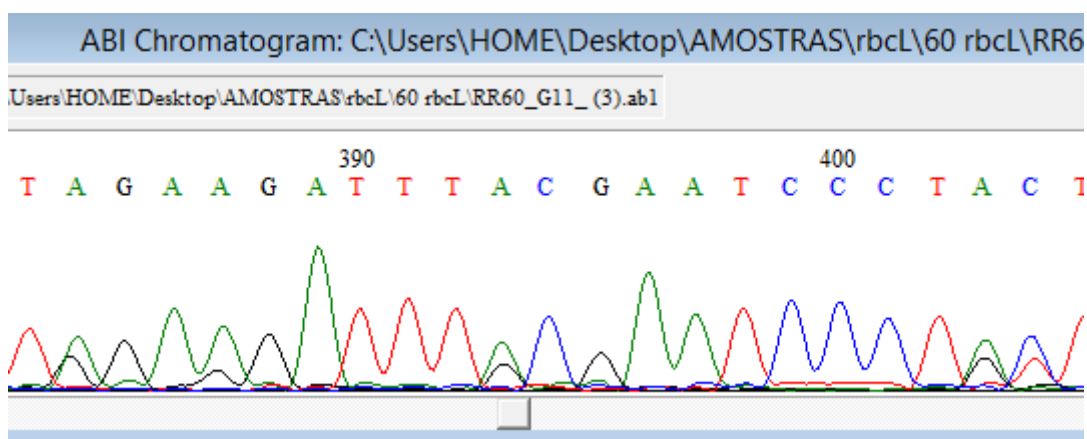


Figura 3 – Eletroferograma da amostra 60, gene *rbcL*, representando picos duplos e indicação de mistura de espécie.

Saslis-Lagoudakis e colaboradores (2015) também identificaram uma sequência quimérica, em que vários picos duplos são observados no DNA, mostrando algumas bases que são características de *E. arvense* e algumas de *E. palustre*, indicando uma mistura de amostras.

4.2.3 Utilização dos genes *matK* e *rbcL* como código de barras de DNA

A combinação de *matK* e *rbcL* foi proposta por como sendo a ideal como padrão para plantas terrestres, pois essa representa uma combinação entre universalidade, qualidade de sequência, discriminação e custo (CAMERON et al., 2009). Entretanto, nem sempre duas ou mais regiões faz-se necessária, como pôde ser observado neste estudo. Como também relatado por Palhares (2015b), em que ficou demonstrado que nem sempre duas regiões são necessárias, por conseguinte, é possível padronizar qual sequência é mais eficiente para cada espécie, diminuindo assim o número de *locus* em estudo e consequentemente tornando a análise ainda mais rápida e menos custosa.

Não houve divergências quanto a identificação das amostras comerciais entre os dois genes em estudo. Ambos os genes foram suficientemente discriminatórios para a identificação molecular das espécies *P. glomerata* e *P. paniculata*. Entretanto, como o gene *matK* apresentou problema no sequenciamento devido uma região de repetição de bases (Figura 4). Assim, o gene *rbcL* parece mais adequado para ser utilizado com *barcode* nesse gênero. Problemas com o sequenciamento do gene *matK* também foram relatados no estudo de Zhang e colaboradores (2015) e Tripathi e colaboradores (2013).

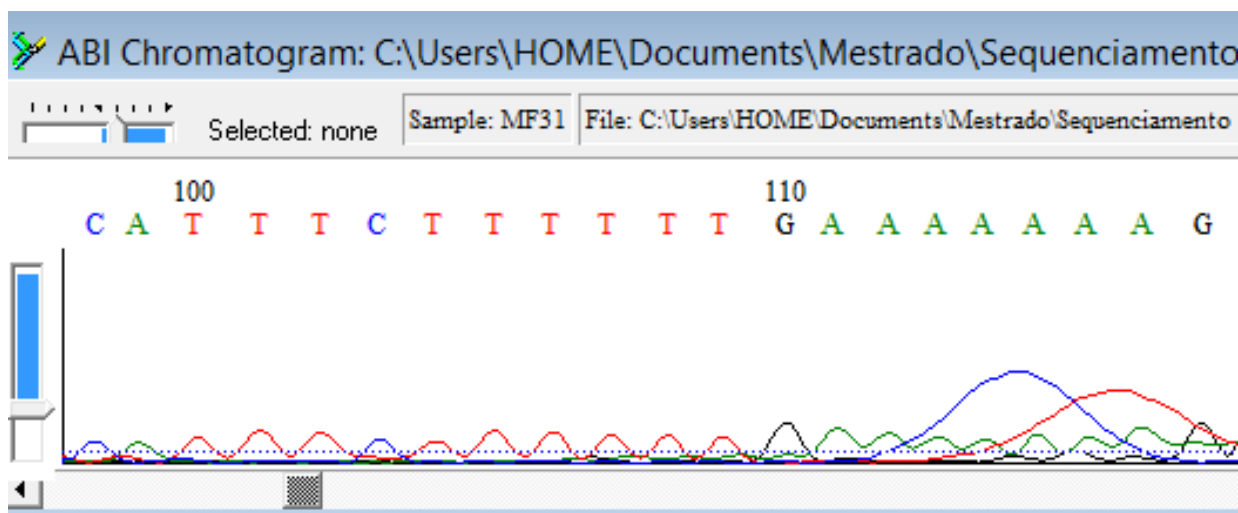


Figura 4 – Eletroferograma da amostra comercializada 31, do gene *matK*, demonstrando repetição de bases.

Cada gene se mostra de maneira diferente de acordo com a espécie. Nem sempre um gene ideal para uma espécie será o adequado para uma outra. Como pode ser notado por Januário (2014), em que o gene *rbcL* não foi tão eficiente para as espécies utilizadas em seu estudo, mesmo sendo uma das regiões mais utilizadas no *barcode* e uma das mais bem caracterizadas no Genbank, (NEWMASTER et al., 2006).

O gene *rbcL* apresentou boa qualidade no sequenciamento e mostrou-se muito eficiente na identificação das espécies estudadas neste trabalho. Ele pôde distinguir todas as amostras analisadas, porém uma das amostras não apresentou amplificação para o gene *rbcL*, somente para o gene *matK*. Tripathi e colaboradores (2013) também obtiveram sucesso em seu trabalho, gerando sequências do *locus* *rbcL* de alta qualidade. Esse gene é considerado um dos *locus* de código de barras mais eficiente, embora não seja uma região muito variável, é frequentemente utilizado em combinações multilocus para discriminação de espécies.

4.3 Importância do BOLD

As amostras referências foram depositadas no BOLD e contribuíram muito para o crescimento do banco de dados visto que ainda não haviam depósitos para as duas espécies em estudo, para os genes *matK* e *rbcL*. Durante o depósito das sequências no banco de dados, percebemos que o gene *matK* apresentou códon de parada. Isso já foi relatado em outros estudos, como de Barthet et al. (2015) e Barthet & Hilu (2007). A rápida taxa de substituição e códons de parada prematuros no gene *matK*, levou os pesquisadores a sugerirem que este gene pode não ser funcional em algumas plantas, sendo considerado, portanto, como um pseudogene (HIDALGO et

al., 2004; JANKOWIAK et al., 2004). Pseudogene é o gene que não forma uma proteína funcional. Existem cerca de 121.000 depósitos de sequência do gene matK no GenBank. Destes, 3.094 são pseudogenes, sendo que desses pseudogenes depositados, cerca de 82% (2.523) são da família Orchidaceae (Harrison et al., 2005; Zheng et al., 2007).

É importante que sejam desenvolvidos estudos para identificação de plantas e que estes colaborem para o desenvolvimento da técnica de identificação molecular. Quanto a esta técnica, não existe um consenso a respeito de quais *loci* utilizar, embora exista uma proposta de combinação. Porém, de acordo com diversos estudos, várias combinações podem ser feitas e o resultado vai depender da espécie em estudo. Cada espécie apresenta uma melhor caracterização utilizando um determinado *locus* ou combinação (KRESS et al., 2005; HOLLINGSWORTH et al., 2009).

A aplicabilidade desta técnica vem sendo avaliada justamente por essa questão (FU et al., 2011; GUO et al., 2011). Vários problemas são relatados devido à complexidade na evolução das espécies, entre elas aloploidia, misturas de reprodução sexual e assexuada, introgressão, hibridação interespecífica, entre outros. Por isso, faz-se necessário o estudo de cada espécie individualmente, para que assim seja analisada a melhor combinação de loci a ser utilizada (SPOONER, 2009; ZUO et al., 2011).

Atualmente, a tecnologia de DNA código de barras é a melhor tecnologia para identificação de espécies (Han et al., 2016). Como pode ser observado neste estudo, essa técnica pode ser utilizada para o controle de qualidade de plantas medicinais, garantindo assim maior segurança e eficácia para o paciente que está utilizando.

Portanto, a metodologia mostrou-se aplicável para o Ginseng-brasileiro uma vez que possibilitou a identificação das plantas comerciais e discerniu as discrepâncias existentes.

5 CONCLUSÃO

O método de identificação por *barcoding* possibilitou a identificação, em nível de espécie, das amostras comercializadas como Ginseng brasileiro obtidas do mercado brasileiro, como *P. glomerata* e *P. paniculata*. Algumas amostras não foram identificadas como nenhuma das duas espécies.

Tanto as amostras referências, quanto comercializadas, foram caracterizadas e algumas amostras não foram identificadas de acordo com o que indicava no rótulo.

Todas as amostras foram analisadas utilizando extração de DNA, reação de PCR e sequenciamento, porém em algumas amostras não foi possível realizar o sequenciamento, visto que não amplificaram na reação de PCR.

As amostras referências foram depositadas no BOLD e mostraram-se 100% idênticas entre si, não apresentando variabilidade genética intraespecífica para os genes *matK* e *rbcL*.

Esse método mostrou-se eficaz e aplicável uma vez que possibilitou a identificação dos espécimes e apresentou diferenças genéticas significativas entre as espécies em estudo, tanto para o gene *matK* quanto para o *rbcL*. Portanto, a metodologia mostrou-se oportuna para o Ginseng-brasileiro uma vez que possibilitou a identificação das plantas comerciais e o confronto com o que estava relatado no rótulo.

6 REFERÊNCIAS

ANVISA. Agência Nacional de Vigilância Sanitária (Brasil) (2008) Instrução Normativa nº 5 de 11 de Dezembro de 2008: Lista de medicamentos fitoterápicos de registro simplificado. Disponível em: <http://bvsms.saude.gov.br/bvs/saudelegis/anvisa/2005/int0005_11_12_2008.html> . Acesso em: 26 set. 2016.

AZEREDO, A. M. L. de. O Código de Barras da Vida baseado no DNA “Barcoding of Life”: Considerações e Perspectivas. Universidade Estadual de Campinas-UNICAMP: 2005. Disponível em: < www.cggee.org.br/atividades/redirect/1749 >. Acesso em: 26 set. 2016.

BALICK, M. J. & COX, P. A. **Plants, people and culture**. New York: Scientific American Library. 1997.

BARRETT, R. D. H.; HEBERT, P. D. N. Identifying spiders through DNA barcodes. **Can J Zool**, v.83, p. 481–491, 2005.

BARTHET, M. M.; HILU, K. W. Expression of matk: functional and evolutionary implications. **American Journal of Botany**, v. 94, n. 8, p. 1402–1412, 2007.

BARTHET, M. M.; MOUKARZEL, K.; SMITH, K. N.; PATEL, J.; HILU, K. W. Alternative translation initiation codons for the plastid maturase MatK: unraveling the pseudogene misconception in the Orchidaceae. **BMC Evolutionary Biology**, v. 15, n. 210, p. 12862-015-0491-1, 2015.

BENNETT, R. N. & WALLSGROVE, R. M. Secondary metabolites in plant defence mechanisms. **New Phytologist** 127, 617–633 doi: 10.1111/j.1469-8137.1994.tb02968.x (1994).

BRANDAO, M.; COSENZA, G. P.; PEREIRA, F. L.; VASCONCELOS, A. S.; FAGG, C. W. Changes in the trade in native medicinal plants in Brazilian public markets. **Environ Monit Assess**, v.185, p. 7013–7023, 2013.

BRASIL. RDC nº 26, de 13 de maio de 2014. **Dispõe sobre o registro de medicamentos fitoterápicos e o registro e a notificação de produtos tradicionais fitoterápicos**. Disponível em: www.anvisa.gov.br. Acesso em: 27 set. 2016.

BRASIL. RDC nº 48, de 16 de março de 2004. **Dispõe sobre o registro de medicamentos fitoterápicos**. Disponível em: www.anvisa.gov.br. Acesso em: 27 set. 2016.

BRUNI, I.; DE MATTIA, F.; GALIMBERTI, A.; GALASSO, G.; BANFI, E.; CASIRAGHI, A.; et al. Identification of poisonous plants by DNA barcoding approach. **Int J Legal Med**, v. 124, p.595–603, 2010.

CALIXTO, J. B. Efficacy, safety, quality control, marketing and regulatory guidelines for herbal medicines (phytotherapeutic agents). **Brazilian J. Med.Biol. Res.** 33: 179-89, 2000.

CAMERON, K. M.; CHASE, M. W.; WHITTEN, W. M.; KORES, P. J.; JARRELL, D. C.; ALBERT, V. A.; YUKAWA, T.; HILLS, H. G.; CBOL Plant Working Group. A DNA barcode for land plants. **Proc. Natl. Acad. Sci. USA**, v.106, n.31, p. 12794–12797, 2009.

CAÑIGUERAL, S.; DELLACASSA, E.; BANDONI, A. L. Plantas Medicinales y Fitoterapia: ¿indicadores de dependencia o factores de desarrollo? **Acta Farm Bonaer**, v. 22, n. 3, p. 265-278. 2003.

CARVALHO, A. C. B.; PERFEITO, J. P. S.; COSTA E SILVA, L. V.; RAMALHO, L. S.; MARQUES, R. F. O.; et al. Regulation of herbal medicines in Brazil: advances and perspectives. **Braz J Pharm Sci**, v. 47, n. 3, p. 467-473. 2011.

CBOL Plant Working Group: A DNA barcode for land plants. **Proc Natl Acad Sci. USA**, v. 106, p.12794-12797, 2009.

CHASE, M. W.; COWAN, R. S.; HOLLINGSWORTH, P. M.; BERG, C.; MADRINAN, S.; PETERSEN, G.; SEBERG, O.; JORGENSEN, T.; CAMERO, K. M.; CARINE, M.; PEDERSEN, N.; HEDDERSON, T. A. J.; CONRAD, F.; SALAZAR, G. A.; RICHARDSON, J.E.; HOLLINGSWORTH, M. L. ; BARRACLOUGH, T. G.; KELLY, L.

; WILKINSON, M. A proposal for a standardized protocol to barcode all land plants. **Taxon**, v.2, n.2, p. 295– 299, 2007.

CHASE, M. W.; SALAMIN, N.; WILKINSON, M.; DUNWELL, J. M.; KESANAKURTHI R. P.; HAIDAR, N.; SAVOLAINEN V. Land plants and DNA barcodes: short-term and long-term goals. **Philos Trans R Soc Lond B Biol Sci**, v. 360, p.1889-1895, 2005.

CHEN, S.; YAO, H.; HAN, J.; LIU, C.; SONG, J.; SHI, L.; et al. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. **PLoS One**, 2010.

CHINA PLANT BOL GROUP; LIA, D. Z.; GAOA, L M.; LIA, H. T.; WANGA, H.; GEA, X. J.; LIUA, J. Q.; CHENA, Z. D.; ZHOUA, S. L.; CHENA, S. L.; YANGA, J. B.; FUA, C. X.; ZENGA, C. X.; YANA, H. F.; ZHUA, Y. Z.; SUNA, Y. S.; CHENA, S. Y.; ZHAOA, L.; WANGA, K.; YANGA, T.; DUANA, G. W. Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. **PNAS**, v. 108, n. 49, p. 19641–19646, 2011.

DANTA, B. F.; ARAGÃO, C. A.; ARAÚJO-JUNIOR, J. P.; RODRIGUES, J. D.; CAVARIANI, C.; NAKAGAWA, J. Padronização da metodologia do rt-PCR utilizado para identificação do mRNA da α -amilase em sementes de milho. **Revista Brasileira de Sementes**, vol. 24, nº 2, p.113-117, 2002.

DE BOER, H. J.; OUARGHIDI, A.; MARTIN, G.; ABBAD, A. & KOOL, A. DNA Barcoding Reveals Limited Accuracy of Identifications Based on Folk Taxonomy. **PLoS ONE**, v.10, 2014.

DE LUCA, V.; SALIM, V.; ATSUMI, S. M.; Y. U. F. Mining the biodiversity of plants: a revolution in the making. **Science**, v. 336, p. 1658–1661, 2012.

ERICKSON, D. L.; SPOUGE, J.; RESCH, A.; WEIGT, L. A.; KRESS, J. W. DNA barcoding in land plants: Developing standards to quantify and maximize success. **Taxon**, 57:1304–1316, 2008.

ERNST, E. Risks of herbal medicinal products. **Pharmacoepidemiology and Drug Safety**, v. 13, p. 767–771, 2004.

FARNSWORTH, N. R. Testando plantas para novos remédios. **Biodiversidade**. Rio de Janeiro, Editora Nova Fronteira, p. 107-125, 1997.

FAZEKAS, A. J. et al. Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. **PLoS ONE**, v. 3, 2008.

FEBRAFARMA. FEDERAÇÃO BRASILEIRA DA INDÚSTRIA FARMACÊUTICA. Disponível em: <<http://www.febrafarma.org.br>>. Acesso em: 20 abr. 2009.

FORD, C. S.; AYRES, K. L.; TOOMEY, N.; HAIDER, N.; STAHL, J. V. A.; KELLY, L. J.; WIKSTROM, N.; HOLLINGSWORTH, P. M.; DUFF, R. J.; HOOT, S. B.; COWAN, R. S.; CHASE, M. W.; WILKINSON, M. J. Selection of candidate coding DNA Barcoding regions for use on land plants. **Botanical Journal of the Linnean Society**, v.159, p.1-11, 2009.

FRACKMAN, S.; KOBBS, G.; SIMPSON, D.; STORTS, D. Betaine and DMSO: Enhancing Agents for PCR. **Promega Notes**, n. 65, p. 27, 1998.

FU YM, JIANG WM, FU CX. Identification of species within Tetrastigma (Miq.) Planch.(Vitaceae) based on DNA barcoding techniques. **J Syst Evol**, v. 49, p. 237–245, 2011.

GILBERT, N. Regulations: Herbal medicine rule book. **Nature**, v. 480, p. 98-99, 2011.

GUERRA, P. M.; NODARI, O. R. Biodiversidade: aspectos biológicos, geográficos, legais e éticos. **Farmacognosia:da planta ao medicamento**. 3. ed. Porto Alegre: UFRGS; Florianópolis: UFSC, p.15, 2001.

GUO, X.; SIMMONS, M. P.; BUT, P. P. H.; SHWA, P. C.; WANG, R. J. Application of DNA barcodes in Hedyotis L. (Spermacoceae, Rubiaceae). **J Syst Evol**, v. 49, p. 203–212, 2011.

HALL, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. **Nucleic Acids Symposium Series: Oxford Journals**, v. 41, p. 95-98, 1999.

HAMILTON, A.C. Medicinal plants, conservation and livelihoods. **Biodiversity and Conservation**, v.13, p. 1477-1517. 2004.

HAN, J.; PANG, X.; LIAO, B.; YAO, H.; SONG, J.; CHEN, S. An authenticity survey of herbal medicines from markets in China using DNA barcoding. **Nature -Scientific Reports**, v. 6, 2016.

HARRISON, P.; ZHENG, D.; ZHANG, Z.; CARRIERO, N.; GERSTEIN, M. Transcribed processed pseudogenes in the human genome: an intermediate form of expressed retrosequence lacking protein-coding ability. **Nuc Acids Res**, v. 33, p. 2374–83, 2005.

HARVEY, A. Strategies for discovering drugs from previously unexplored natural products. **Drug Discovery Today**, v. 5, n. 7, p. 294-300, 2000.

HEBERT, P. D. N.; GREGORY, T. R. The promise of DNA Barcoding for taxonomy. **Systematic Botany**, v.54, p.852-859, 2005.

HEBERT, P. D. N.; PENTON, E. H.; BURNS, J. M.; JANZEN, D.H.; HALLWACHS, W.: Ten species in one: DNA barcoding reveals cryptic species in neotropical skipper butterfly *Astraptes fulgerator*. **Proc Natl Acad Sci USA**, v. 101, p.14812-14817, 2004.

HEBERT, P. D. N.; STOECKLE, M. Y.; ZEMLAK, T. S.; FRANCIS, C. M.: Identification of birds through DNA barcodes. **PLoS Bio**, ed. 2, 2004.

HEBERT, P.D.; RATNASINGHAM, S.; DEWAARD, J.R. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. **Proc. Biol. Sci.**, v. 7, p. S96-9, 2003.

HILU, K. W.; e LIANG, H. The matK gene: sequence variation and application in plant systematics. **American Journal of Botany**, v.84, n.6, p. 830– 839, 1997.

HOLLINGSWORTH, M. L. Selecting barcoding loci for plants: Evaluation of seven candidate loci with species-level sampling in three divergent groups of land plants. **Mol Ecol Res**, v. 9, p. 439–457, 2009.

HOLLINGSWORTH, P. M.; FORREST, L. L.; SPOUGE, J. L.; HAJIBABAEI, M.; RATNASINGHAM S. et al. A DNA barcode for land plants. **Proc Natl Acad Sc USA**, v. 106, p. 12794–12797, 2009.

HOSTETTMANN, K.; MARSTON, A. Saponins-chemistry and pharmacology of natural products. **Cambridge University**, 1995.

JANUARIO, B. B. Aplicação de códigos de barras de DNA (DNA barcoding) na identificação das espécies de *Senna Mill.* (*Fabaceae*) e *Casearia Jacq.* (*Salicaceae*) para estudos de variabilidade genética. **Dissertação de mestrado**. Faculdade de Ciências Farmacêuticas, Araraquara, 2014.

JENSEN, M. A.; FUKUSHIMA, M.; DAVIS, R. W. DMSO and Betaine Greatly Improve Amplification of GC-Rich Constructs in De Novo Synthesis. **Plos One**, v. 5, ed. 6, p. 11024, 2010.

JORDAN, S. A., CUNNINGHAM, D. G. & MARLES, R. J. Assessment of herbal medicinal products: Challenges, and opportunities to increase the knowledge base for safety assessment. **Toxicology and Applied Pharmacology**, v. 243, p. 198–216, 2010.

KENNEDY, J. Herb and supplement use in the US adult population. **Clinical Therapeutics**, v.27, p. 1847–1858, 2005.

KOOL, A. Molecular Identification of Commercialized Medicinal Plants in Southern Morocco. **PLoS ONE**, v. 7, 2012.

KOSMANN, C. Código de Barras (DNA Barcode) de dípteros de interesse forense. 2009. 61 f. **Dissertação (Mestrado)**. Universidade Federal do Paraná, Curitiba. 2009.

KRESS, W. J.; ERICKSON D. L. A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. **PLoS One**, v. 2, p. 508, 2007.

KRESS, W. J.; WURDACK, K. J.; ZIMMER, E. A.; WEIGT, L. A.; JANZEN, D. H. Use of DNA barcodes to identify flowering plants. **Proc Natl Acad Sci USA**, v. 102, p. 8369–8374, 2005.

KRESS, W.; ERICKSON, D. L. A Two-Locus Global DNA Barcode for Land Plants: The Coding *rbcL* Gene Complements the Non-Coding *trnH-psbA* Spacer Region. **Plos One**, ed. 6, 2007.

LAHAYE, R. DNA barcoding the floras of biodiversity hotspots. **Proc Natl Acad Sci USA**, v. 105, p. 2923–2928, 2008.

LEDFOORD, H. Botanical identities: DNA barcoding for plants comes a step closer. **Nature** 451:616, 2008.

LI, J. W. H.; VEDERAS, J. C. Drug discovery and natural products: End of an Era or Endless Frontier. **Science**, v. 325, n. 5937, p. 161-165. 2009.

LI, S. Chemical markers for the quality control of herbal medicines: an overview. **Chinese Medicine**, v. 3, p. 7, 2008.

LORENZI, H. & MATOS, F.J.A. **Plantas Medicinais no Brasil: nativas e exóticas**. 2ª ed. Nova Odessa, Instituto Plantarum, 2008.

MASOUD, S. A.; JOHNSON, L. B.; WHITE, F. F. The sequence within two primers influences the optimum concentration of dimethyl sulfoxide in the PCR. **PCR Methods Applications**, v.2, p.89-90,1992.

MINISTÉRIO DA SAÚDE. **Programa Nacional de Plantas Medicinais e Fitoterápicos**. Disponível em: <
http://bvsms.saude.gov.br/bvs/publicacoes/politica_nacional_fitoterapicos.pdf >.
 Acesso em 15 set 2016.

NEWMASER, S. G.; FAZEKAS, A. J.; RAGHUPATHY, S. DNA barcoding in land plants: evaluation of rbcL in a multigene tiered approach. **Canadian Journal of Botany**. v. 84, p. 335–341, 2006.

NEWMASER, S.; RAGUPATHY, S.; JANOVEC, J. A botanical renaissance: state-of-the-art DNA barcoding facilitates an Automated Identification Technology system for plants. **Int J Comput Appl T**, v. 35, n.1, 2009.

NICOLOSO, F. T.; ERIG, A. C.; MARTINS, C.;F; RUSSOWSKI, D. Micropropagação do ginseng brasileiro [*Pfaffia glomerata* (Spreng.) Pedersen]. **Ver Bras Plant Med**,v. 3, n.2, p. 11-8, 2001.

NODARI, R. O.; GUERRA, M. P. Biodiversidade: aspectos biológicos, geográficos, legais e éticos. **Farmacognosia da planta ao medicamento**. Porto Alegre, Ed UFRGS/ UFSC, 1999.

OLIVEIRA, F. de. *Pfaffia paniculata* (Martius) Kuntze – o ginseng-brasileiro. **Revista Brasileira de Farmacognosia**, v.1, p.86-92, 1986.

PALHARES, R. M. Aplicação da tecnologia de DNA Barcode em espécies vegetais aprovadas pela ANVISA e comercializadas no Brasil e elaboração de metodologia para certificação de fitoterápicos. **Tese de doutorado**, Universidade Federal de Minas Gerais, Belo Horizonte, 2015b.

PALHARES, R. M.; DRUMMOND, M. G.; BRASIL, B. S. A. F.; COSENZA, G. P.; BRANDÃO, M. G. L.; OLIVEIRA, G. Medicinal Plants Recommended by the World Health Organization: DNA Barcode Identification Associated with Chemical Analyses Guarantees Their Quality. **Plos One**, v. 10, n. 5, 2015.

PALHARES, R. M.; DRUMMOND, M. G.; BRASIL, B. S. A. F.; KRETTLI, A.U.; BRANDÃO, M. G. L.; OLIVEIRA, G. The use of an integrated molecular-, chemical – and biological-based approach for promoting the better use and conservation of medicinal species: A case study of Brazilian quinins. **Journal of Ethnopharmacology**, v. 155, p. 815-822, 2014.

PENNISI, E. Taxonomy. Wanted: A barcode for plants. **Science**, v. 318, p.190–191, 2007.

PRANCE, G. T. Floristic inventory of the tropics: where do we stand? **Missouri Bot. Gard.**, v.64, p. 559-684, 1977.

RALSER, M.; QUERFURTH, R.; WARNATZ, H. J.; LEHRACH, H.; YASPO, M. L.; KROBITSCH, S. An efficient and economic enhancer mix for PCR. **Science Direct**, v. 347, n. 3, p. 747–751, 2006.

RATNASINGHAM, S.; HEBERT, P. D. N. BOLD: The Barcode of Life Data System (www.barcodinglife.org). **Mol. Ecol. Notes**, v. 7, p.355-364, 2007.

SAHOO, N.; CHOUDHURY, K.; MANCHIKANTI, P. Manufacturing of biodrugs: Need for harmonization in regulatory standards. **Bio Drugs**, v. 23, n. 4 p. 217-229. 2009.

SASLIS-LAGOUDAKIS, C. H.; BRUUN-LUND, S.; IWANYCKI, N. E.; SEBERG, O.; PETERSEN, G.; JÄGER, A. K.; RØNSTED, N. Identification of common horsetail (*Equisetum arvense* L.; Equisetaceae) using Thin Layer Chromatography versus DNA barcoding. **Nature - Scientific Reports**, v. 5, 2015.

SEBERG, O.; PETERSEN, G. How many loci does it take to DNA barcode a crocus? **PLoS ONE**, v. 4, 2009.

SHENKEL, E. P.; GOSMANN, G.; PETROVICK, P. R. **Produtos de origem vegetal e o desenvolvimento de medicamentos. Farmacognosia: da planta ao medicamento**. 3 ed. Porto Alegre/Florianópolis: Editora da Universidade UFRGS / Editora da UFSC, cap. 15, p. 301-332, 2001.

SHIOBARA, Y.; INOUE, S-S.; NISHIGUCHI, T.; OISHI, Y.; NISHIMOTO, N.; OLIVEIRA, F.; AKISUE, G.; AKISUE, M. K.; HASHIMOTO, G. Pfaffane-type nortriterpenoids from *Pfaffia pulverulenta*. **Phytochemistry**, v. 31, p. 1737-1740, 1992.

SHU, Y. Recent natural products based drug development: A pharmaceutical industry perspective. **Journal of Natural Products**, v. 61, p. 1053-1071, 1998.

SINGH, H. K.; PARVEEN, I.; RAGHUVANSHI, S.; BABBAR, S. B. Loci recommended as universal barcode for plants on the basis of floristic studies may not work with congeneric species as exemplified by DNA barcoding of *Dendrobium* species. **BMC Res. Notes**, v.5, n. 42, 2012.

SPOONER DM. DNA barcoding will frequently fail in complicated groups: an example in wild potatoes. **Am J Bot**, v. 96, p. 1177–1189, 2009.

STOECKLE, M. Y.; GAMBLE, C. C.; KIRPEKAR, R.; YOUNG, G.; AHMED, S.; LITTLE, D. P. Commercial Teas Highlight Plant DNA Barcode Identification. **Scientific Reports**. v.1, n.42, 2011.

STOECKLE, M.; P.E. WAGGONER & J.H. AUSUBEL. **Barcoding Life Illustrated. Goals**, 2005.

TAKIUCHI, E.; MÉDICI, K. C.; ALFIERI, A. F.; ALFIERI, A. A. Otimização da reação em cadeia pela polimerase (Semi Nested-PCR) para a detecção do herpesvírus

bovino tipo 1 em fragmentos de órgãos fetais e em sêmen de bovinos naturalmente infectados. **Semina: Ciências Agrárias**, v. 24, n. 1, p. 43-56, 2003.

TECHEN, N., CROCKETT, S. L., KHAN, I. A. & SCHEFFLE, B. E. Authentication of Medicinal Plants Using Molecular Biology Techniques to Compliment Conventional Methods. **Current Medicinal Chemistry** 11, 1391–1401, 2004

TRIPATHI, A. M.; TYAGI, A.; KUMAR, A.; SINGH, A.; SINGH, S.; CHAUDHARY, L. B.; ROY, S. The Internal Transcribed Spacer (ITS) Region and trnhHpsbA Are Suitable Candidate Loci for DNA Barcoding of Tropical Tree Species of India. **PLOS ONE**, v. 8, 2013.

VAN BREEMEN, R. B., FONG, H. H. S. & FARNSWORTH, N. R. Ensuring the safety of botanical dietary supplements. **The American Journal of Clinical Nutrition** 87, 509S–513S, 2008.

VIDAL, M.R.R; VIDAL, W.N. Flórmula de Viçosa. I. Chenopodiaceae e Amaranthaceae. **Rev. Ceres.**, v. 14, p. 46-79, 1967.

VIEIRA, R. A. Validação científica de plantas medicinais como fator catalisador no desenvolvimento da indústria farmacêutica nacional. **Rev Saúde Amb**, v. 2, n. 1/2, p.57-64. 2001.

WACHTEL-GALOR, S.; BENZIE, I. F. F. Herbal Medicine: An Introduction to Its History, Usage, Regulation, Current Trends, and Research Needs. In: Benzie IFF, Wachtel-Galor S, editors. **Herbal Medicine: Biomolecular and Clinical Aspects**. 2 ed., 2011.

WHO (World Health Organization). **Quality control methods for medicinal plants materials**. Geneva, p. 114, 1998.

WHO (World Health Organization). **Monographs on selected medicinal plants**. v. 1, p. 295, 1999.

WHO (World Health Organization). **Monographs on selected medicinal plants**. v. 2, p. 375, 2003.

WHO (World Health Organization). **Medicinal plants—guidelines to promote patient safety and plant conservation for a U\$ 60 billion industry**. 2004.

WHO (World Health Organization). **Monographs on selected medicinal plants** v. 3, p. 390, 2007.

WHO (World Health Organization). **Monographs on selected medicinal plants** v. 4, p. 456, 2009.

WHO (World Health Organization). **Monographs on selected medicinal plants**, v. 5. Geneva, p. 452, 2010.

WHO (World Health Organization). **The World Traditional Medicines Situation, in Traditional medicines: Global Situation, Issues and Challenges**. v. 3, p.1–14, 2011.

WHO (World Health Organization). **Bulletin of the World Health Organization. Regulatory situation of herbal medicines**. 1998.

ZHANG, J. Q.; MENG, S. Y.; WEN, J.; RAO, G. Y. DNA Barcoding of *Rhodiola* (Crassulaceae): A Case Study on a Group of Recently Diversified Medicinal Plants from the Qinghai-Tibetan Plateau. **PLOS ONE**, v 10, n.3, 2015.

ZHENG, D.; FRANKISH, A.; BAERTSCH, R.; KAPRANOV, P.; REYMOND, A.; CHOO S. W. Pseudogenes in the ENCODE regions: Consensus annotation, analysis of transcription, and evolution. **Genome Res**, v.17, p.839–51, 2007.

ZUO, Y. J.; CHEN, Z. J.; KONDO, K.; FUNAMOTO, T.; WEN, J.; ZHOU, S. L. DNA barcoding of *Panax* species. **Planta Med**, v. 77, p.182–187, 2011.

ANEXO 1

Amostras vendidas como ginseng brasileiro de diversos locais do país.

Número	Nome comercial na embalagem	Lote	Local de compra	Tipo	Parte
1	<i>Pfaffia</i> (ginseng brasileiro)	12015	internet	Pó	X
2	Ginseng (<i>Pfafia Paniculata</i>)	52468	internet	Pó	X
3	Ginseng Brasileiro	110/14	internet	Pó	X
5	Ginseng	X	Casa das Ervas - Vila Rubim	Cápsula	X
6	Ginseng	X	Natura Minas - Vila Velha	Fragmento	Folhas, flores e talos
7	Ginseng (<i>Pfafia paniculata</i>)	2	Mundo Verde - Vila Velha	Pó	Raiz
8	Ginseng	X	Pequena Selva - Vila Rubim	Pó	X
9	Ginseng	X	Casa das Ervas - Vila Rubim	Pó	X
10	Ginseng	X	Mercearia Rodrigues - Vila Rubim	Pó	X
11	Ginseng (<i>Pfafia paniculata</i>)	FAFI001	Natura Minas - Vila Velha	Pó	Raiz
12	Ginseng	X	Pharma Ervas - Vila Rubim	Pó	X
13	<i>Panax</i>	X	Guarapari	Pó	X
14	Ginseng	X	Araraquara-SP	Pó	X
15	Ginseng	X	Manaus- AM	Pó	X
16	Ginseng (<i>Pfaffia</i> spp)	32	Rio de Janeiro-RJ	Fragmento	Raiz
17	Ginseng	X	Rio de Janeiro-RJ	Fragmento	Raiz
18	Ginseng (<i>Pfafia paniculata</i>)	X	Rio de Janeiro-RJ	Pó	X
19	Ginseng em Pó	X	Barra da Tijuca-RJ	Pó	X
20	Ginseng Pó Varejo	X	Rio de Janeiro-RJ	Pó	X
21	Ginseng Chinês (<i>Panax ginseng</i>)	1220	Ribeirão Preto-SP	Pó	X
22	Ginseng	X	Ribeirão Preto-SP	Fragmento	Raiz
23	Ginseng Nacional (<i>Pfaffia paniculata</i>)	1224	Ribeirão Preto-SP	Pó	X
24	Ginseng		Ribeirão Preto-SP	Pó	X
25	<i>Pfafia</i> (<i>Pfaffia glomerata</i>)	2581	Ribeirão Preto-SP	Pó	X

Amostras vendidas como ginseng brasileiro de diversos locais do país.

Número	Nome comercial na embalagem	Lote	Local de compra	Tipo	Parte
26	Ginseng	X	Belo Horizonte-MG	Cápsula	X
27	Ginseng	X	Belo Horizonte-MG	Cápsula	X
28	Ginseng	X	Belo Horizonte-MG	Cápsula	X
29	Ginseng	X	Belo Horizonte-MG	Pó	X
30	Ginseng	X	Belo Horizonte-MG	Pó	X
31	Ginseng	X	Belo Horizonte-MG	Pó	X
32	Ginseng Brasileiro (<i>Pfaffia paniculata</i>)	X	Belo Horizonte-MG	Pó	X
33	Ginseng	X	Belo Horizonte-MG	Pó	X
34	Ginseng	X	Belo Horizonte-MG	Pó	X
35	Ginseng	X	Belo Horizonte-MG	Pó	X
36	Ginseng (<i>Pfaffia paniculata</i> Mart.)	X	Belo Horizonte-MG	Pó	Raiz
37	Ginseng	X	Belo Horizonte-MG	Pó	X
38	Ginseng	X	Belo Horizonte-MG	Pó	X
39	Ginseng (<i>Pfaffia</i>)	X	Belo Horizonte-MG	Pó	X
40	Ginseng	X	Belo Horizonte-MG	Pó	X
41	Ginseng Brasileiro (<i>Pfaffia Paniculata</i>)	X	Belo Horizonte-MG	Pó	X
42	Korean Ginseng Tea gold	X	Belo Horizonte-MG	Pó	X
43	Ginseng	X	Belo Horizonte-MG	Pó	X
44	Ginseng	X	Belo Horizonte-MG	Fragmento	Folhas, flores e talos
45	Ginseng	X	Belo Horizonte-MG	Pó	X
46	Ginseng	X	Belo Horizonte-MG	Pó	X
49	Ginseng	x	Guaçuí-ES	Pó	X
50	Ginseng (<i>Pffafia Paniculata</i>)	X	Guarapari	Pó	X
51	Ginseng em pó	X	São Paulo	Pó	X
52	Ginseng pó	X	São Paulo	Pó	X
53	Ginseng em pó	X	São Paulo	Pó	X
54	Diversos	X	São Paulo	Pó	X
55	Ginseng em pó KG	X	São Paulo	Pó	X
56	Ginseng em pó	X	São Paulo	Pó	X

Amostras vendidas como ginseng brasileiro de diversos locais do país.

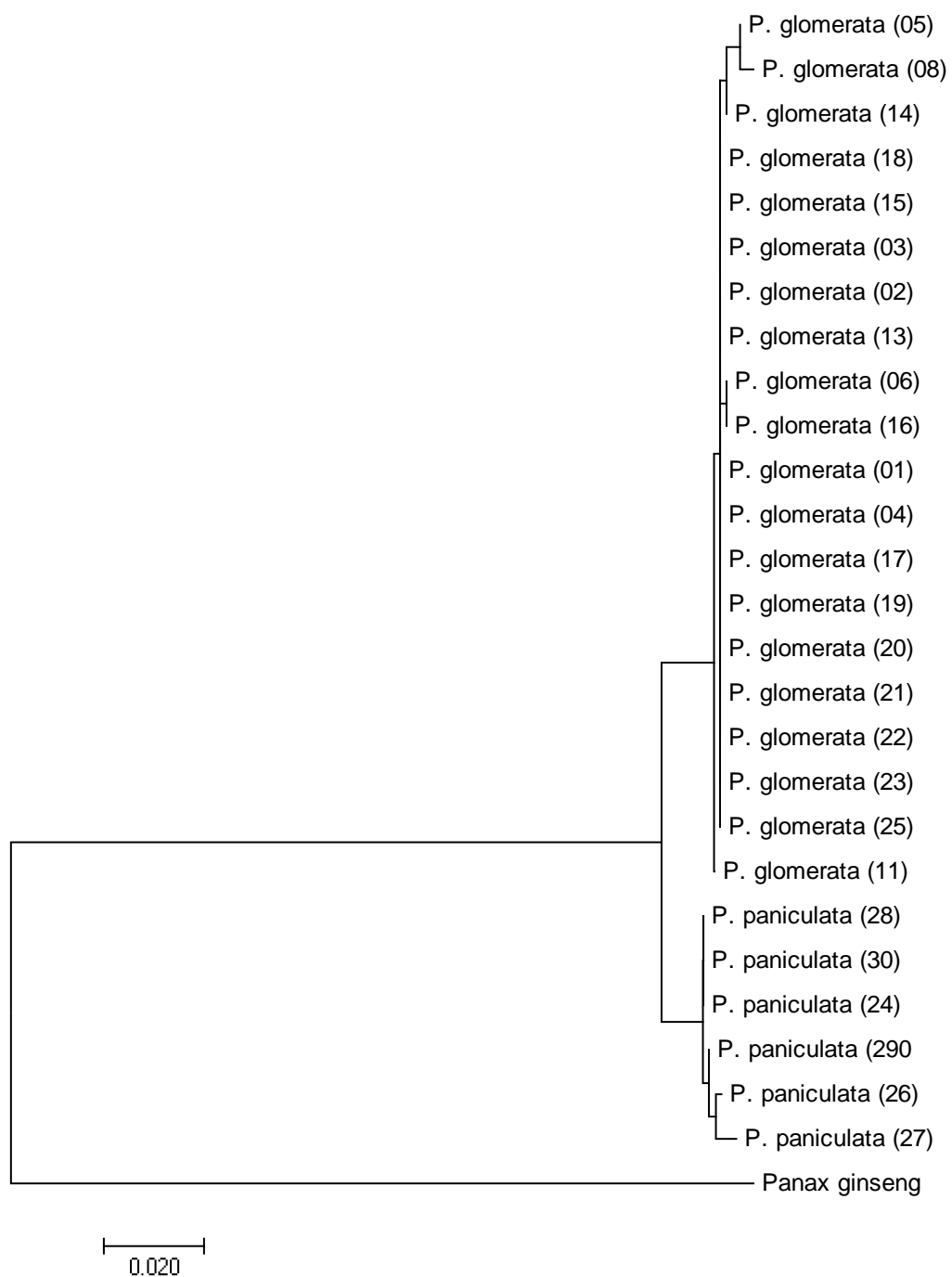
Número	Nome comercial na embalagem	Lote	Local de compra	Tipo	Parte
57	Ging Seng	X	São Paulo	Pó	X
58	Ginseng raiz	X	São Paulo	Raiz	Raiz
59	Ginseng em pó 12	X	São Paulo	Pó	X
60	Ginseng pó granel	X	São Paulo	Pó	X
61	Ginseng pó	X	São Paulo	Pó	X
62	Ginseng- 209	X	São Paulo	Pó	X
65	Ginseng granel	X	São Paulo	Pó	X

Legenda: X – não disponibiliza a informação no rótulo.

ANEXO 2

Árvores filogenéticas das amostras referências juntamente a uma amostra do gênero *Panax*, retirada do BOLD, para os genes *matK* e *rbcL*. Utilizou-se o programa MEGA7, método NJ.

Árvore filogenética gene *matK*



Árvore filogenética gene rbcL

